ONTARIO. ENGLISH. INTERMEDIATE ~~SENIOR~~ DIVISIONS. SUPPORT DOCUMENTS.
1979. OHEC/373.23609713/059DE/IE
3668

Ministry of Education
Ontario

Hon. Bette Stephenson, M.D., Minister
Dr. H. K. Fisher, Deputy Minister

Curriculum Ideas for Teachers

1979

English
Intermediate
Division

CURRICULUM RESOURCES
OISE LIBRARY

Evaluation and the English Program

A Support Document for
English, Intermediate Division, 1977

COMP. STOR.

The Ministry of Education wishes to acknowledge the contributions of the following educators in the preparation of this document:

Writing Team
P. J. Evans, Ontario Institute for Studies in Education, Sudbury Office (Chairman)
Marvle Bell, Teacher, Sudbury Board of Education
Cyril Gounden, Teacher, Niagara South Board of Education
Paul Legge, Consultant, London and Middlesex County Roman Catholic Separate School Board
Marion Marsh, Ontario Institute for Studies in Education, Sudbury Office

Validators
Jean R. Baxter, Lennox and Addington County Board of Education
Patricia Beatty, Metropolitan Separate School Board
Warren Campbell, Scarborough Board of Education
Sheila E. Chalat, Lakehead Board of Education
Micki Clemens, Halton County Board of Education
Diane Cole, Board of Education for the City of Hamilton
Jim Doris, Research and Evaluation Branch, Ministry of Education
Helen Dunlop, Research and Evaluation Branch, Ministry of Education
John Fitchett, Ottawa Board of Education
Patricia Hughes, Elementary Education Branch, Ministry of Education
Kenn Johnson, Central Ontario Regional Office, Ministry of Education
Bill Lipischak, Western Ontario Regional Office, Ministry of Education
Margaret Wolchak, Etobicoke Board of Education

Ministry of Education Co-ordinators
Jerry George, Elementary Education Branch
Catherine Michalski, Policy, Liaison and Legislation Branch

Contents

Evaluation: Purposes, Methods, and Reporting	3
Introduction	3
Diagnostic Evaluation	4
Formative Evaluation	4
Summative Evaluation – By the Teacher	5
Responding to Literature	6
Introduction	6
Student Activities	6
Evaluating Students' Response to Literature	6
Bibliography	8
Evaluating Student Reading	8
Introduction	8
Diagnosis of Reading Problems	8
Formative Evaluation: Alternatives in the Evaluation of Reading	10
Summative Evaluation	14
Bibliography	15
Evaluating Writing	15
Introduction	15
Characteristics of a Writing-Evaluation Scheme	16
Mode, Audience, Purpose, Register	18
The Student Writing Folder	18
Specific Criteria	18
Holistic, Multiple Marking	18
Diagnostic Procedures	19
Writing Workshop	19
Bibliography	19
Language Study and Evaluation	20
Introduction	20
Language Fundamentals and Writing: The Individual	20
Language Fundamentals and Writing: The Group	20
Specific Testing of Language Competence	22
Explorations in Language	23
Expansion of Vocabulary	24
Oral Language	25
Bibliography	25
Evaluation: School, Program, System	26
Introduction	26
Essentials for a Systematic Approach	26
An Evaluation Program Within the School	26
A Checklist for Teachers, Principals, Consultants, and Administrators	27
Testing and Sampling in a School System	28
Reporting	31
Bibliography	32
General	32
Journals	33
Objectives in English	33
Reading	33
Writing	33

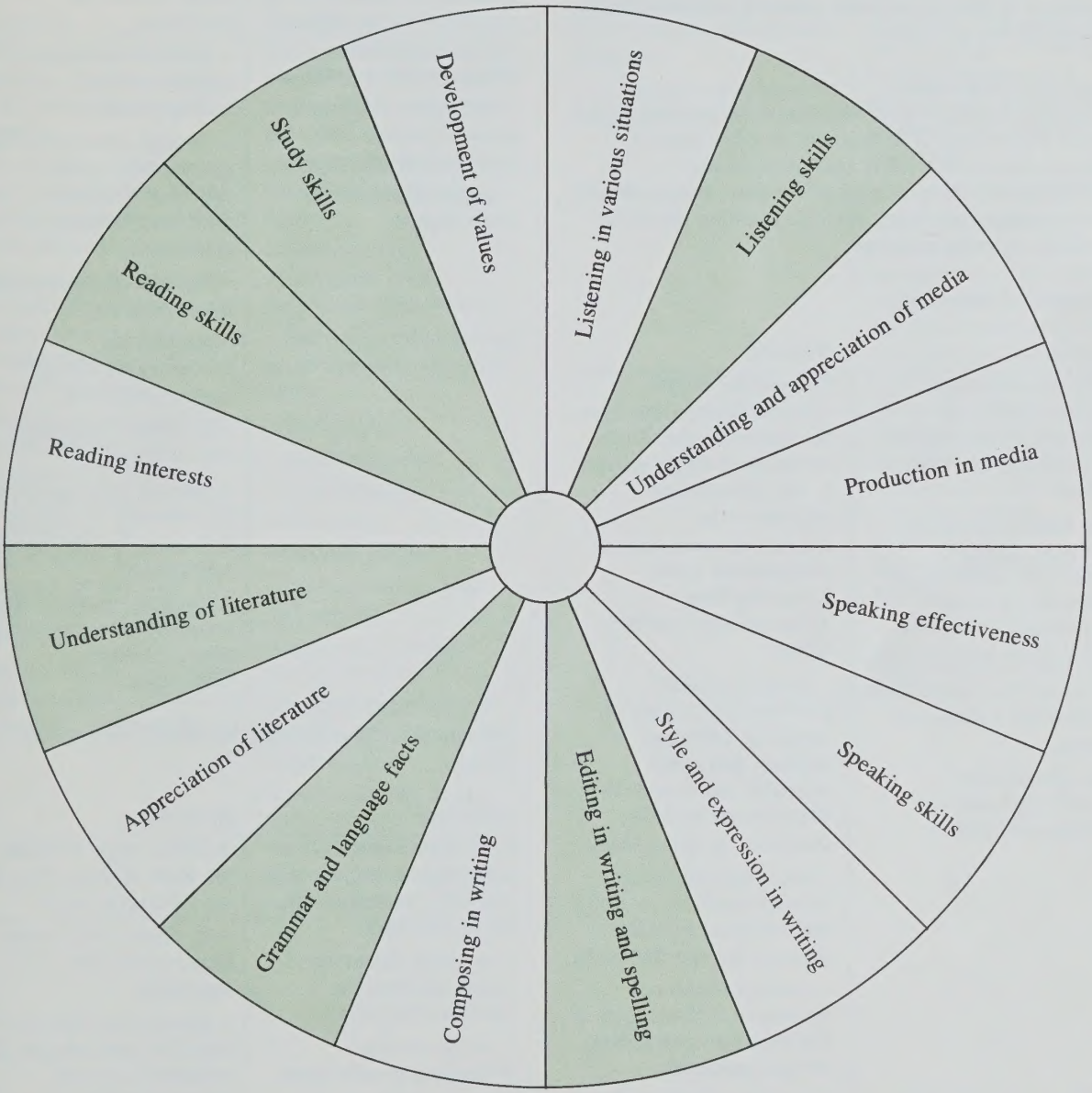
Evaluation: Purposes, Methods, and Reporting

Introduction

Because evaluation is an essential component of the teaching/learning process, the evaluation techniques used by the teacher must be carefully chosen or constructed. They should relate to learning objectives and clearly indicate student achievement and program effectiveness.

Evaluation should be diagnostic and formative, providing information on what is necessary for further learning; it should be summative, providing information about student achievement and program effectiveness and promoting parental and community awareness in these areas; and it should aim to develop in students habits of self-evaluation.

As well, the results of evaluation must be reported in ways that inspire confidence in the validity and reliability of the methods of evaluation and that reflect the diversity of English. This diversity is well illustrated by the following diagram:



Some of the Many Facets of English
The shaded sections are those partially covered by standardized English tests.[1]

[1]*Common Sense and the Teaching of English,* Report of the Task Force on Measurement and Evaluation in the Study of English (Champaign, Ill.: National Council of Teachers of English, 1975), p. vi.

This diagram shows the range of language skills and indicates those areas that can be partially tested by existing, standardized English tests. There is obviously a danger in using standardized tests to make judgements about a student's achievement in this field. There is also a need for a reporting system that will recognize the huge range and integrated nature of the subject.

The following charts are meant to summarize and clarify the different types of evaluation necessary in the English program. The methodologies or strategies suitable for each type of evaluation are also listed. However, the charts are by no means exhaustive. The terms used in these charts are more fully explained and illustrated in their appropriate contexts in the remaining sections of this document.

Diagnostic Evaluation

Diagnostic testing is most valuable at the beginning of a school year or unit of instruction, or when there is evidence that an individual student is having difficulties. At the beginning of the year, testing should be comprehensive. Later, individual testing should be limited to specific concerns.

Diagnostic Evaluation

Purposes:	*Methods:*
– To assess the skills, interests, abilities, difficulties, and level of achievement of a class or group; – to identify the skills, interests, abilities, difficulties, and level of achievement of an individual; – to make decisions about the program modifications suitable for a particular group; – to identify the underlying causes of learning difficulties.	The teacher might: –review information from the previous year. Such information might include: 1. test information (whether from teacher-designed tests or standardized tests); 2. reading lists; 3. the student's writing folder; – observe student performance in class in speaking, listening, reading, and group activities and record the observations by using checklists or rating scales; – use diagnostic tests (teacher-made or carefully selected from existing sources) for specific skills; – conduct individual interviews with students to discover their perceptions of their problems; – employ attitude scales or attitude sorts to discover the attitudes of individuals in a class.

Formative Evaluation

Formative evaluation provides continuous information on how the program is functioning and what students are achieving. This information forms the basis for further planning for, and teaching of, individuals and groups of students. The concepts of formative, diagnostic, and summative evaluation are elaborated further in *English, Intermediate Division,* pages 106-7.

Formative Evaluation – By Students of Their Own Progress

Purposes:	*Methods:*
– To assist students in evaluating their own progress; – to contribute to students' awareness of the objectives of the course and of competent language performance; – to develop students' sense of responsibility for their own development; – to provide feedback from students.	Students might evaluate their own work in the following ways: – by observation, using rating scales and checklists; – through anecdotal reports (written self-appraisals); – through interviews with the teacher – oral self-assessment of progress, strengths, weaknesses; – through their questions for the teacher; – through the requirements of their student contracts; – by proofreading their own material; – through their selection of materials for writing folders or for the teacher's assessment; – by observations of their own performances through tape recordings or videotapes.

Formative Evaluation – By Member(s) of the Peer Group

Purposes:	*Methods:*
– To provide the student with other responses to, and other audiences for, his or her work; – to create situations in which students can compare their work; – to allow more evaluation to take place; – to provide further records to be considered in summative evaluation.	Students might evaluate the work of their peers in the following ways: – by observation, using rating scales and checklists; – through the writing of anecdotal and comparative comments on each student's work; – by proofreading each other's essays; – through discussions about each other's work; – through questions about each other's work; – through student-made quizzes;

	– through marking clinics; – through the use of a panel of evaluators (especially for oral presentations). *Note:* Peer evaluation may be conducted not only in a class, but also between classes, levels, grades, small groups, or partners.

Formative Evaluation – By the Teacher

Purposes:	*Methods:*
– To measure individual and class growth in skills in all areas of language learning; – to measure individual and class growth in effort, attitudes, and appreciation of the subject; – to provide feedback to the individual, class, and teacher on progress towards the objectives of the unit or course; – to indicate which skills or attitudes are at a satisfactory level and which need improvement; – to evaluate the effectiveness of a program, in terms of either its content or its methods; – to provide records that will form part of a summative evaluation.	The teacher might make use of the following tools as part of the evaluation process: – observation, using checklists and rating scales; – interviews or conferences with individuals and groups; – the writing of anecdotal reports on each student's work; – flow charts; – questionnaires; – class-climate inventories; – attitude scales; – tape recordings and videotapes; – analysis of student writing folders; – teacher-made tests (chiefly criterion-referenced). *Note:* Using the same technique at different times during the course gives a basis for comparison.

Summative Evaluation – By the Teacher
Summative evaluation occurs at the end of a unit, course, or program and is used to determine student achievement and program effectiveness.

Purposes:	*Methods:*
– To measure relative student achievement; – to give or withhold credit; – to report to parents, principal, and students; – to monitor the overall performance of students in English; – to measure the effect of program change.	Teachers might use the following tools in their evaluation of their students' achievement in English: – holistic scoring of student writing; – analysis of a carefully drawn sample of student writing, referenced to error counts, criteria, and measures of complexity and style; – teacher-designed tests; – matrix sampling of student performance based on objectives (In matrix sampling all students do not answer all questions. A sample set of questions is answered by a sample group of students, who remain anonymous. The overall effect of the school's programs on the overall academic achievement of its students can thus be assessed.); – criterion-referenced testing based on an item pool.

Responding to Literature

Introduction

In the Intermediate Division, the experience of literature is far more important than knowledge *about* literature. Thus, the development of a positive student *attitude* and *response* should be the central aim of the literature program at this level.

This aim requires the use of an "experience-based" approach, involving the affective as well as the cognitive aspects of learning. In the course of the "response to literature" approach, students should actively engage in listening, speaking, viewing, acting, reading, and writing activities, individually and in groups. Some of these activities should be observed, recorded, and evaluated in a variety of ways. The following are some student activities that can be evaluated.

Student Activities

Generally, student activities in literature may involve:

- small-group discussion and oral presentations through which students explore literature;
- dramatizing scenes from literature;
- reading and taping scenes from literature;
- rewriting literature in contemporary language and/or style;
- writing notes on characters and themes;
- comparing themes in literature;
- listening to short stories, poems, and plays;
- reading aloud stories, poems, and plays;
- sustained silent reading in class.

Evaluating Students' Response to Literature

Evaluation should be an open process between teacher and student, in which every student should be encouraged to trace his or her own progress. The following suggestions may help to activate this open evaluation process:

1. Both students and teacher keep a log book to record the student's progress.
2. The student writes a personal evaluation, tracing his or her growth over a period of time.
3. The teacher and student meet to discuss the student's development.
4. The teacher makes clear his or her expectations to the students.
5. The teacher provides opportunities for students to evaluate their own work.
6. The teacher keeps a record of student participation in class by observing small groups and individual students.

Assessment strategies may include checklists, inventories, questionnaires, and rating scales, and may be used both by the teacher and by students. Teachers and students must understand both the criteria and the methods used to evaluate student activities.

In order to assess their own work and the work of their peers accurately, students may find it helpful to use simply constructed measuring instruments. One such evaluation instrument is a rating scale that outlines the criteria and the relative value of each one (see Figure 3). The models presented in Figures 1 to 5 may be modified to serve the individual needs of teachers and students.

Figure 1: Observing and Evaluating Responses
Put a minus sign (−) in the appropriate box in the "A" column against each activity in which the student participates. If the learner shows improved participation in subsequent evaluations, indicate this with a plus sign (+).

Student _____ Date _____

Type of Participation	A	B	C	D
1. Listening – in class/group	☐	☐	☐	☐
2. Speaking – in class/group	☐	☐	☐	☐
3. Reading – silently/aloud	☐	☐	☐	☐
4. Writing – notes, research, etc.	☐	☐	☐	☐
5. Other – viewing, rehearsing, etc.	☐	☐	☐	☐

Comments _____

_____

_____

_____

*Figure 2: Student Reading Record**

Name: _____ Grade: _____

Date	Book read	Pages	Comments
___	_____	___	_____
___	_____	___	_____
___	_____	___	_____
___	_____	___	_____
___	_____	___	_____
___	_____	___	_____
___	_____	___	_____
___	_____	___	_____
___	_____	___	_____
___	_____	___	_____

Comments:

* This form is to be completed by the student.

*Figure 3: Evaluation of Student Activities**
Type of student activity:

1. Oral presentations to the class by individuals or groups (oral responses to literature, sharing book reports/reviews, discussing literary experiences, themes, characters)
2. Dramatized performances by groups of students
3. Interviews, seminars, and panel presentations

Criteria and scale for evaluating each of the above:

Was the presentation

1. effective?	1	2	3	4	5
2. relevant?	1	2	3	4	5
3. clear and convincing?	1	2	3	4	5
4. useful in promoting sharing and learning?	1	2	3	4	5
	low				high

Ratings:
5 – Excellent 4 – Very Good 3 – Good
2 – Satisfactory 1 – Poor

* This model is based on the evaluation of activities in literature and drama and may be suitable for both student and teacher evaluation. Writing assignments based on literature may be evaluated as suggested in the "Evaluating Writing" section of this resource document.

Figure 4: Peer Evaluation of an Oral Presentation

	Score
1. Was the presentation well planned?	_____
2. Did you find the presentation informative?	_____
3. Did the speaker show a good grasp of the subject?	_____
4. Did the speaker articulate his or her responses to questions clearly?	_____
5. On the whole, did the presentation meet the standard you expected?	_____
Total:	_____

Ratings:
5 – Excellent 4 – Very Good 3 – Good
2 – Satisfactory 1 – Poor

*Figure 5: Criteria and Scale to Measure the Growth of a Group**

1. How clear are the goals of the group?

1	2	3	4	5
no goals	confused and uncertain	average	mostly clear	very clear

2. How much trust and openness are there in the group?

1	2	3	4	5
none	little	average	consid-erable	remarkable

3. How much co-operation exists in the group?

1	2	3	4	5
none	little	average	consid-erable	remarkable

4. How are the contributions of individuals received and used by the whole group?

1	2	3	4	5
dissenters silent	dissenters' contributions rejected	average	well-used and en-couraged	effective

* This scale may be completed by individual members of a group. The data can then be shared among members of the group in order to solve any difficulties impeding the group's progress.

The student's appreciation and enjoyment of literature is always a primary consideration. The deductive method of teaching literature – which results in an examination of such elements as plot, character, theme, or setting – tends to emphasize an analysis of the structural components of literature rather than an exploration of such emotions as love, hate, joy, triumph, or guilt. These emotions are abundant in literature and are important elements in the growth of the student's appreciation of literature. In short, the learner should be allowed to relate his or her own experiences to literature.

Bibliography

Adland, D. E. *Group Drama*. London: Longmans, 1964.

Andrews, Larry. "Responses to Literature". *English Journal 66* (February 1977): 60-69.

Ayliffe, Ted; Laird, Ken; Lynch, John; and Pennington, Gary. *Curriculum Design '73: A Resource Booklet*. Toronto: Ontario Secondary Education Commission of OSSTF, 1973.

Riley, Roberts, and Schaffer, Eugene. "Testing Without Tears". *English Journal 66* (March 1977): 65.

Evaluating Student Reading

Introduction

The aims of the reading program, as set out in *English, Intermediate Division, 1977,* pages 11-12, are to enable each student to:

1. acquire the interests and attitudes of a maturing reader;
2. extend and refine personal use of reading/reasoning/language processes;
3. extend and consolidate vocabulary and word skills;
4. develop control over personal reading rate.

It is the responsibility of the teacher to measure individual and class progress towards these goals in order to understand what students can and do read, and what skills they have mastered or must work towards. Self-assessment by students should also be encouraged. Such evaluation is essential to planning an effective reading program for both the class and individual students.

The methods and techniques that follow can be used to *record* systematically the teacher's observations of the students' reading behaviour, and the students' perceptions of their own reading behaviour, in order to obtain over a period of time some *reasonably valid* information as a basis for evaluation, feedback, and remediation.

The emphasis here is on involving the student in self-evaluation of his or her reading progress. The student must become motivated to read and to be aware of and able to express the difficulties he or she experiences in reading. The teacher's role as "motivator" is emphasized.

Some system of maintaining a record or file for each student is implied in the following strategies. The system should be convenient for the teacher to use and easily understood by the student. Such records will be valuable to the teacher in planning the reading program and for reporting to students and parents.

Diagnosis of Reading Problems

Several suggestions are given in the section on "Formative Evaluation" for techniques to assess student interests and general abilities at the beginning of the school year. This section focuses on diagnosing some of the more severe reading problems. The following checklist and methods should not be used for all students in the class; they may, however, be useful for dealing with students who are obviously having severe or puzzling problems. Furthermore, they are not intended as sensitive, precise instruments to aid in the clinical diagnosis of the underlying reading problem. They are merely suggestions to help the classroom teacher understand the nature of the problem more clearly. The following statement suggests a good approach to the diagnosis of reading problems:

> A person's failure to learn to read can seldom, if ever, be explained by a single cause. Few people fail to learn to read because of a single visual or auditory deficiency. When you try to help one of your secondary students identify why he has had trouble, you will likely discover emotional or behavioural problems which overlay the original causes. For some, not-reading has become a comfortable escape

and, for the moment, a quite acceptable life-pattern. Knowing that, in fact, not-reading can seldom remain acceptable in our world, you have a major task in changing their attitude while you exercise your prerogative to teach them differently. Our justification for asking you to attempt to identify the predominant causes for each individual's failure is that, as Dr. Ruth Strang used to say, "If we are going to stretch out a helping hand to him, we had better make sure we are standing on the same corner of the block as he is." This identification ought to be a factor in our method of remediation and in our plan to help him understand himself as a student.

From *Developmental Reading: A Resource Booklet* (Toronto: OSSTF, 1972), p. 10.

A Diagnostic Checklist

Note: This checklist is not meant to be comprehensive – the range of potential problems is too great. It may, however, be suggestive of some possible problems.

Note: The teacher puts a small checkmark in the box opposite the observation made.

Student's Name: _____

Teacher's Name: _____

	1st observation	2nd observation	3rd observation
Visual problems			
1. Frowns or squints			
2. Rubs eyes frequently			
3. Points with finger			
4. Loses place from line to line frequently			
Auditory problems			
5. Has difficulty hearing (following directions) when speaker faces away from him or her			
6. Has difficulty in hearing when there is background noise (records playing, students talking)			
7. Speaks too loudly			
Speech problems			
8. Has a speech defect			
9. Speaks very seldom			
Perceptual and other problems			
10. Reverses letters in writing/reading letters and numbers			
11. Has poor memory for written materials or instructions			
12. Has poor handwriting (cramped, poorly spaced)			
13. Has poor motor co-ordination ("clumsy")			
14. Has short attention span			
15. Is hyperactive (or constantly taps toe, pencils)			
16. Is "emotional", "irritable", "aggressive", "impulsive"			
17. Any other unusual behaviour noted by teacher			

What can the teacher do about such problems?

Students with visual or auditory problems should be referred to eye specialists or medical doctors.

Often there is no resource person (reading consultant, psychometrist, guidance teacher) to aid the teacher who has students with speech, perceptual, or other related problems. If this is the case, the teacher will have to proceed carefully, in close consultation with the student, to remedy any problems.

Miscue Analysis

Miscue analysis is an effective diagnostic and research instrument that has been developed to help teachers assist students in improving their reading skills.The premise of the test is that the "miscues" a reader makes when reading a passage are not necessarily signs of reading weaknesses. Often they indicate that the student *is* understanding or "getting meaning" from what he or she is reading.

The formal "Reading Miscue Inventory", although not difficult to understand and implement, is too lengthy to reproduce in full in this document. The interested reader should refer to: Y. M. Goodman and C. L. Burke, *Reading Miscue Inventory: Procedure for Diagnosis and Evaluation* (New York: Macmillan, 1972).

The following summary of the principles of miscue analysis is taken from *Communications – Language Guideline, Intermediate,* published by the Board of Education for the Borough of East York.

Principles That Are Basic to the Concept of Miscue Analysis

– In reading, meaning is always central.
– What we need to measure is the reader's ability to keep his focus successfully on meaning. That means we are primarily concerned with the quality of his miscues.
– It follows then that rather than a mere meticulous counting of miscues we are concerned with their effect on meaning.
– The reader needs to learn to use meaning and grammatical structure to detect when he has made a miscue and to learn to eliminate those which will hinder his receiving the message of the writer.
– The teacher uses the insights into the reader's experience with print to devise strategy lessons to strengthen his ability to keep his focus on the meaning.

Questions the Teacher Should Ask About the Reader's Miscues

1. Does the reader rely on the sound-letter similarity to the exclusion of concern with gaining meaning?
2. Are the miscues which the reader corrects the ones which he should have corrected?
3. Does the reader translate somewhat consistently certain words or phrases into his own dialect?
4. Does the student habitually associate the same words with each other? Are the associations disruptive to meaning?
5. Does the student indicate through his miscues that he does not understand certain concepts or ideas which are presented by the author?
6. Is the reader able to understand certain concepts and ideas even if he mispronounces words or phrases related to the concept?
7. To what extent does the reader succeed in interrelating the graphophonic with the semantic and syntactic cueing system as he reads?

Questions to Be Asked About Words and Sentences Containing Miscues

Words

1. How much do the two words *look* alike? (graphic similarity)
2. How much do the two words *sound* alike? (sound similarity)
3. Is the grammatical function of the reader's word the same as the grammatical function of the text word? (grammatical or *syntactical* similarity)

Sentences

4. Is the sentence involving the miscues *syntactically* (grammatically) acceptable?
5. Is the sentence involving the miscues *semantically* acceptable? (meaning)
6. Meaning change: Is there a change in meaning involved in the sentence?
7. Do corrections by the reader make the sentence semantically acceptable?

A Note on Standardized Tests

The objectives of the reading program, as outlined in the Intermediate Division guideline for English, are deeper and more far-reaching than can be measured by any multiple-choice standardized test consisting of rather factual, short pieces of reading material. Such standardized tests can measure only relative achievement in a rather small number of reading skills. Further information about the use and limitations of standardized tests can be found in the Intermediate Division English guideline, beginning on page 108.

Formative Evaluation: Alternatives in the Evaluation of Reading

1. Questionnaires

The questionnaire is an excellent method of discovering students' interests in, and attitudes towards, reading. Easily constructed by the teacher and quickly completed by students, the questionnaire can provide information about the individual student's reading development and can suggest new directions or modifications for the reading program. Questions should be brief, clear, and easily understood.

At the beginning of a school term, the teacher may wish to make a quick check of where students are in their reading development. A "reading interest questionnaire" will acquaint the teacher with the student and his or her perception of his or her own reading habits. The teacher should follow up the questionnaire with a personal interview in which the student has the opportunity to respond with greater thought to the teacher's questions. Meanwhile the completed questionnaires may suggest some themes for the reading program, class or individual projects, or subjects for class discussion.

Reading Interest Questionnaire
I want to learn a little more about you and your interests. This is not a test; there are no right or wrong answers. Please answer as completely as possible.

Name: _____ Date: _____

1. Which subject do you like best at school?

_____

2. What games or sports do you like to play?

_____

3. What television programs to you like to watch?

_____

4. What are your interests or hobbies outside of school?

_____

5. How often do you read the newspaper? _____
What section do you like best? _____
6. Do you read any magazines? _____
Which ones do you like? _____
7. Do you read comic books? _____
Which are your favourites? _____
8. Do you like to read books?
If so, what kind?
a) short stories _____
b) adventure stories _____
c) mystery stories _____
d) science fiction _____
e) stories about the lives of people _____
f) books that explain how to make or do things
_____
g) other (please specify) _____
9. Do you ever get books from either the library at school or the public library? _____
10. How do you feel about reading? Are there any special reasons you feel this way? _____

_____

_____

Note: Further examples of reading interest questionnaires may be found in: J. M. Barmore and P. Morse, "Developing Lifelong Readers in the Middle Schools", pages 57-61; and D. N. Fader and E. B. McNeil, *Hooked on Books,* pages 191-99.

Later on in the year, when students have gained confidence in themselves and in their teacher, more imaginative and probing questionnaires may elicit a good deal of information about progress in reading. These questionnaires may have quite different aims – the exploration of students' developing interests, their reactions to the reading program, or their frustrations and problems in reading.

The following questionnaire represents one effective way of finding out (a) how readily students become involved in what they read, and (b) what capacities they have for reading various kinds of literary works.

Involvement Measure Questionnaire[1]
There are no right or wrong answers to the following questions. Your answers should be thoughtful and honest.
1. While reading a short story or a novel, have you ever thought of yourself as one of the characters? _____
If so, give an example. That is, give the character's name and/or the title of the book or story. _____

_____

2. Have you ever met someone who you felt was like a character in a play or movie that you had seen or in a story you had read? _____

_____

3. a) When you read – particularly when you read a story or a play – do you ever make a mental picture of the setting? _____

_____

b) If so, how often is that mental picture like some actual place you have seen? Often? _____
Sometimes? _____ Seldom? _____ Never? _____
4. How often do you feel that people you see in TV dramas or in the movies are real people and not actors playing their roles? Often? _____ Sometimes? _____
Seldom? _____ Never? _____

2. Teacher-Student Interviews
During class time devoted to personal reading, the teacher will be able to schedule personal interviews with each student to discuss his or her reading. Such interviews will give the teacher many opportunities to inquire about the student's personal reading habits, interests, and any difficulties that he or she may be experiencing.

Initially, questions should be as open as possible to allow the student to report any difficulties of which he or she is aware. Too often, detailed quizzing of the student provides only answers that the student thinks the teacher expects.

Questions may deal with what the student is currently reading, his or her knowledge of books that he or she has read, and what he or she finds easy or difficult to read.

After an interview, the teacher should quickly record insights into the student's reading behaviour, either on a checklist or, with a few anecdotal remarks, in the student's record or file. Later in the year, such brief remarks will reveal whether there has been a change in the student's interests or behaviour.

[1]Adapted from C. R. Cooper and A. C. Purves, *A Guide to Evaluation* (New York: Ginn, 1973).

The teacher should act on the information obtained, suggesting to the student reading material that is both interesting and challenging. To quote from the Intermediate Division guideline for English:

> . . . the teacher must exercise great care to provide materials that lie just at the edge of the student's expanding knowledge, experience and language abilities.

Teacher-student interviews are discussed more fully in the support document to the Intermediate Division English guideline entitled *Reading*.

3. Checklists

A checklist is a useful technique for teachers to use to record their observations of students' reading behaviours. Students' names are listed down the left side of the checklist, and the reading skills or reading program objectives are listed across the top. During the day, as the teacher listens to a student reading orally, observes a student reading silently, or questions a student on what he or she is reading, the teacher can note any difficulties by placing a checkmark in the appropriate square.

When most of the students have been observed at least twice, the teacher can check across the chart to identify individual students' potential weaknesses and down the chart to identify areas where the whole class, or particular groups of students, have exhibited potential weak areas. The teacher then has a basis for tentative grouping of students in the reading program.

The checklist that follows is very general and can be used throughout the day, no matter what subject is being taught. The objectives or skills being observed will change according to the current emphasis in the reading program, and according to the students' mastery of a particular objective. For instance, if the whole class is using dictionaries satisfactorily, the teacher will concentrate his or her attention on another area. Sometimes only two or three skill areas will be observed at one time.

Reading Skills Checklist[2]

Pupils' Names	Vocabulary			Speed		Comprehension			
	recognizes sight words appropriate to level	able to guess meaning of new words in context	uses dictionary/ encyclopaedia	skims	adjusts speed to purpose	identifies main idea	identifies important details	identifies sequence of ideas/time	can interpret what has been read at a literal level
Suzanne L.		✓ ✓		✓				✓	✓ ✓
Kirk B.			✓						✓
William C.	✓ ✓	✓ ✓		✓	✓		✓		

[2]Adapted from a suggested checklist by Margaret Early in *Developmental Reading – A Resource Book* (Toronto: OSSTF, 1972), p. 21.

Checklists can also be used to introduce students to self-evaluation. The aim of the following checklist, which the teacher prepares, is to have students keep a record of the various reading activities in which they engage each day. The information obtained may be kept in a file and is not only encouraging and valuable to students, but may also be helpful to the teacher. The checklist should cover a specific period of time, which should be long enough to reveal changes in quantity or type of reading.

Self-Evaluation Checklist[3]

Material read	Sept. 5	Sept. 6	Sept. 7	Sept. 8	Sept. 9	Sept. 28
Short story from reader						
Stories by other students						
Passage of history text						
Passage of science text						
Library book						
Other (newspaper, magazines, etc.)						
Total						

4. Student Self-Evaluation: Informal Oral-Reading Evaluation

Students should be encouraged to use a tape recorder from time to time to tape their own oral reading. They might choose an amusing incident they have been reading in a book, a short scene or sketch from a play, or a story that they have written themselves. After silently reading the passage a few times in preparation, they read their chosen passage into the tape recorder. They then check their own performance, listening to the tape and reading the passage silently at the same time.

When time permits, the teacher can also evaluate the student's performance, noting particularly such criteria as intonation, fluency, and articulation. Students may be given a second chance to improve their reading of the passage, at which time any improvement should be noted. As a variation, groups of students may want to record a short play, or a passage read chorally.

5. Charting or Graphing of Reading

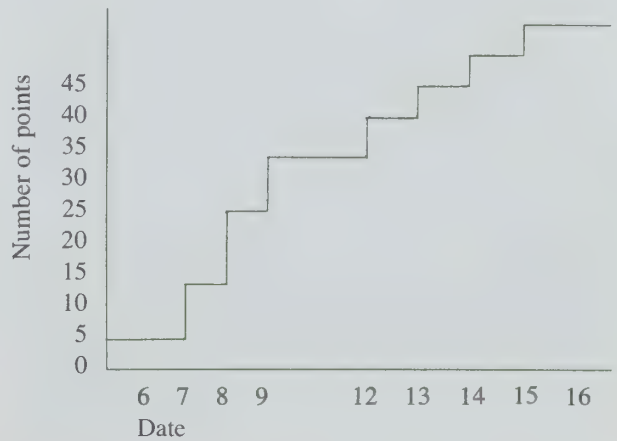
This form of evaluation is especially useful with reluctant readers or slow learners. Two examples are outlined here.[4]

Some students need to see their progress noted concretely and frequently. A chart or graph of the number of pages they have read each week provides them with a sense of accomplishment. A goal should be set beforehand that the student can accomplish; for example: "By the end of the first week, I will have read twenty-five pages." The student then graphs the amount of reading he or she does each week, and keeps this chart on file. (The teacher may want to check this with the student from time to time.) In this way, the student will be able to see reading fluctuations over a lengthy period of time.

As an alternative, instead of counting the pages read, students can allot themselves points for completing a specific reading task. Points might be assigned on the following basis:

– for reading a short story	10 points
– for reading a student-written story	5 points
– for reading a magazine article	5 points

The student should decide on the total points to be gained in one week, two weeks, and three weeks. The gradual accumulation of points would be indicated as shown in the following graph. Individual success is thus noted, day by day and over the weeks.



[3]Adapted from the Sudbury Roman Catholic Separate School Board and the Sudbury Board of Education, *English Curriculum Guide 7-10* (Sudbury: Sudbury Board of Education and Sudbury Roman Catholic Separate School Board, 1974).

[4]Adapted from K. Weber, *Yes They Can!* (Toronto: Methuen, 1974).

6. Book Reports

In preparing book reports, students might be given a practical motive for keeping records on books they have read. For instance, they could write very short book reports (two to three lines plus a personal recommendation or comment) on cards that would be kept in a file in the class. They would sign their names to these reports. Other students could then consult this file when searching for a certain type of book (adventure, humour) to find the ones that were most popular with their classmates.

The teacher could use the file as a record of what books each student has read, and also to discover which books were currently popular in order to recommend similar books that might be appealing to the students. As an alternative to this kind of report, students might prefer to create an ''advertisement'' for a book they particularly enjoyed. These could be posted in the classroom or library.

In their reading discussion groups, students should be encouraged to give oral reports on the books they have enjoyed reading and to answer questions on the book posed by the other students. Such brief presentations could be rated from time to time, either by the teacher or by the other students, on their persuasiveness, clarity, or depth of understanding.

7. Informal Teacher-Made Tests

Teachers may wish to set their own classroom tests to supplement the formative evaluation information they have already collected. In creating their own reading tests – whether of the ''objective'' multiple-choice or fill-in-the-blank type, or the more open-ended ''answer-the-question'' type – teachers should remember the following points:

a) The reading passage chosen should be long enough to stand on its own and should orient the student to its context. Much better results will be obtained if the passage has some intrinsic interest for the students. Humorous stories, advertisements, or magazine articles, for example, would be likely to appeal to most students.

b) The teacher should define and clarify the objectives to be tested before deciding what questions to ask the students. The questions should then relate to these objectives. The objectives should be limited for any one test. For example, the teacher may wish to test student ability to identify the main idea and supporting details in a passage after completing a reading unit concentrating on developing these skills. The teacher should then omit questions that require students to make inferences or reach conclusions, unless the information thus obtained is to be used simply for planning the program and not for evaluating the students.

c) In marking the test, the teacher should concentrate on the individual student's achievement of the objective and should not compare the student's work with that of brighter students in the class.

Asking students to write a summary of an article, a short story, or a chapter from one of the textbooks in science or history is another approach to evaluating skill in reading. If such a method is used, the marking scheme should reflect the reading skills involved, not the writing skills. For instance, if the teacher asked the students to read a short newspaper editorial (500 words) and write a summary of the article in 150 words, the marking scheme might look something like this:

	Low	*Middle*	*High*
a) Stated the main idea of the editorial.	1	2	3
b) Mentioned supporting arguments (five mentioned in article).	1	2	3
c) Maintained the point of view of the writer.	1	2	3
d) Bonus point for staying within the word limit.			1

Student's mark out of 10:

Summative Evaluation

Summative evaluation is described in the Intermediate Division English guideline as ''the periodic (end of term, end of year) total assessment of a student's achievement in relation to his peers and to the course objectives'' (page 107). If even one or two of the suggested formative evaluation techniques have been used, the teacher will have more information on which to judge a student's progress in the reading program. If the teacher has kept some records of the student's reading interests and skills from the beginning of the year through to the last term, it will be easier to determine how much progress the student has made in interest, skill, and range in reading.

The Cloze Procedure for Evaluating Readability

The cloze procedure can be used to determine the appropriateness, in terms of its readability, of a particular book for a particular group of students. The following is a greatly simplified version of this procedure:

1. From the book in question select a passage of at least 500 words of continuous prose that the students have not yet encountered.

2. Type the passage (double-spaced), putting a blank in place of every tenth word, starting with the second sentence. Leave the first and the last sentences intact.

3. Ask the class to fill in the blanks with the missing words.

4. Score the responses, counting only the exact words substituted. (Accept misspellings if you are sure of the word intended.)

The book in question is appropriate for any student who scores at least 44 per cent and not above 70 per cent. (Research indicates that students learn most effectively from books in which they score between 44 per cent and 70 per cent.)

Evaluating the Reading Program
The following is a checklist for the teacher:

Space and Materials
- Does the seating arrangement allow for individual and group teaching?
- Are there books, magazines, and other reading materials available in the classroom?

Scheduling
- Is class time made available each day for reading for pleasure?
- Is enough time allotted for class and individual work on reading skills?

Teacher Planning
- Do I have a definite aim for each lesson?
- Do I have reading materials available at different levels of difficulty and suitable for different interests?
- Do I have varied activities and materials ready to use with individuals, reading groups, or the class?
- Am I sufficiently familiar with the materials the students are reading?
- Am I able to suggest follow-up reading materials when necessary?
- Is each reading activity an integral part of the language program?

Evaluation
- Does evaluation take place throughout the year?
- Are more than one or two different evaluation techniques or strategies used?

Bibliography

Barmore, J. M., and Morse, P. "Developing Lifelong Readers in the Middle Schools". *English Journal* 66 (April 1977): 57-61.

Cooper, C. R., and Purves, A. C. *A Guide to Evaluation*. New York: Ginn, 1973.

Developmental Reading: A Resource Booklet. Toronto: OSSTF, 1972.

Fader, D. N., and McNeil, E. B. *Hooked on Books*. New York: Berkeley Medallion Books, 1966.

Goodman, K. S., ed. *Miscue Analysis*. Urbana, Ill.: ERIC Clearinghouse, 1973.

Goodman, Y. M., and Burke, C. L. *Reading Miscue Inventory: Procedure for Diagnosis and Evaluation*. New York: Macmillan, 1972.

Strang, R.; McCullough, C. M.; and Traxler, A. E. *The Improvement of Reading*. New York: McGraw-Hill, 1967.

Weber, K. *Yes They Can!* Toronto: Methuen, 1974.

Evaluating Writing

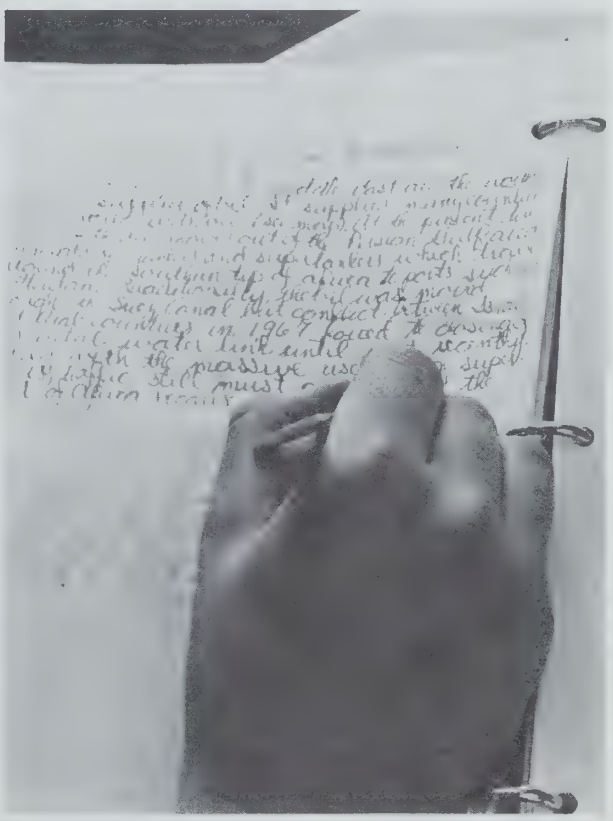
Introduction

> The curriculum . . . inevitably shrinks to the boundaries of evaluation; if your evaluation is narrow and mechanical, this is what your curriculum will be.[1]
>
> Walter Loban

This section provides a series of suggestions for a scheme to evaluate writing that avoids the double-edged danger referred to by Loban. The dangers of narrow writing curricula and narrow, mechanical evaluation schemes have long been with us.

The evaluation of writing is the single most difficult task for the teacher of English, simply because, in evaluating the writing of students, we are assessing much more than their grasp of a program; we are, in the end, evaluating the students themselves.

The three major writing goals and twenty-two learning opportunities set out in the central policy statement of the Intermediate Division English guideline point towards a wide and diverse writing program. This diversity demands that we review our evaluation practices and pay increased attention to the research being done throughout Ontario and elsewhere on the evaluation of student writing.



[1]Quoted in John Dixon, *Growth Through English* (London: Oxford University Press, 1967), p. 92.

Characteristics of a Writing-Evaluation Scheme

A writing-evaluation scheme for a school must be:

– comprehensive	– constructive	– specific
– comprehensible	– informed	– sensitive
– cumulative	– consistent	– manageable
– flexible	– co-operative	– prismatic

A Comprehensive Scheme

Any assessment of students must be based on a wide range of writing activities and a wide sampling of student writing. In the Intermediate Division guideline for English, the policy statement and the section entitled "Development of Writing Abilities" are consistent in establishing the necessity for students to complete writing activities in various modes, for various purposes and audiences, and in different registers. The evaluation scheme should reflect that variety. Similarly, any research into the development of writing abilities will point to the necessity for much writing by students. That quantity of writing, which at first may seem to present an unbearable marking load for the teacher, requires a new approach to marking or monitoring that involves the use of the student writing folder, specific criteria, and holistic, multiple marking.

Finally, an evaluation scheme for writing must consider both prewriting and postwriting activities, in view of the fact that the student's preparation for writing and examination of his or her own completed work and that of his or her peers is crucial to the learning process. As well, the students' writing should be integrated with the total language program.

A Constructive Scheme

The major goals of education in this province require that confidence and a sense of self-worth be considered of equal value with basic skills. Thus, sound evaluation rests on a positive, supportive base. Research findings corroborate the view that constructive evaluation of writing meets the student where he or she is (through careful diagnostic procedures) and takes him or her forward. Consequently, all evaluation schemes must measure progress in student writing and must have built-in formative measures that react specifically and constructively to writing problems. To this end, error charts, specific criteria, and the use of multiple markers at specified times in the program are recommended.

A Specific Scheme

There is no justification for the evaluation of student writing by vague and generally indefinable methods. Students should know what they are expected to do and whether or not they (and everyone else being evaluated) have done it. The development of specific criteria for the evaluation of writing, the maintaining of cumulative, anecdotal records, and the necessity for adhering to the Ministry writing goals are all critical to a valid scheme.

A Comprehensible Scheme

An evaluation and reporting system should show clearly what it is we are teaching in writing, and how these goals are or are not being met in individual student programs. Any evaluation scheme should adequately represent the progress made by a student in writing in terms that evolve from specific criteria, that recognize clearly established objectives, and that involve the participation of the student, his or her peers, and the teacher. The section on reporting will explore these ideas further.

An Informed Scheme

A valid evaluation scheme reflects the range of characteristics that constitute a piece of writing. Teachers should be aware of the importance of prewriting (talking) activities; of indications of maturity in language patterns; of mode, audience, purpose, and register; of the lack of good, standardized test materials in writing; and of the minor significance of grammatical error in the overall pattern of writing development.

The teacher's understanding of how written language works and, more importantly, of how writing skills are acquired is crucial in evaluating a student's work. Major studies in England and the United States have a message for the teacher of writing: When we are better informed about the teaching of writing skills and when we specify clearly what we are doing and how well our students are responding, we will be much better prepared to respond to our students.

A Sensitive Scheme

Evaluation can be highly productive if, before a writing assignment is undertaken, students are made aware of the evaluative criteria to be used and the purpose, mode, audience, and register for the writing assignment. Emphasis must also be placed on motivating the student. The evaluation processes we use must build skills, of course, but they must also build feelings of confidence and self-worth; they must be sensitive to the absolute necessity to separate performance from personality.

A Cumulative Scheme

The evaluation of writing should be an ongoing process. The process should monitor the regular and developing growth of writing skills, and should demonstrate openly that changes have taken place. To this end, the use of the writing folder for each student in the Intermediate Division, and the development of a procedure whereby that folder follows the student from grade to grade, will pay obvious dividends. Samples of student writing (preferably the best, as chosen by the teacher and the student), cumulative and anecdotal comments, and error data – in sum, a portfolio of what the student can do and has done in this specific subject area – will be of immeasurable help to the next teacher responsible for the student's continuing acquisition of writing skills.

A Consistent Scheme

Respectability and validity will accompany any evaluation scheme that is consistent at the school level. Such a scheme will have strong effects on both the Junior and Senior divisions. Its adoption by a division

within a school, and a concomitant presentation of the scheme to the rest of the staff, will produce increased levels of students and staff confidence. A writing program consisting of clearly established criteria, demonstrable student progress, and the element of objectivity afforded by holistic, multiple marking will develop confident teachers of writing, teachers whose grades are knowledgeably assigned.

A Manageable Scheme

If initiated with planning and care, the evaluation scheme presented here will mean: (a) more writing by students, (b) less marking by staff, and (c) more reliable evaluation overall.

Several levels of evaluation (self, peer, teacher), the maintenance of a writing folder and other cumulative records showing the acquisition of skills for which the student is responsible, and a reduction in the number of summative evaluations required make the scheme work. Teachers who have been involved in variations of this process report more time spent in individual and group discussions with students about writing skills and less time in the onerous paper grading that has long characterized the role of the writing teacher.

A Flexible Scheme

A sensitive evaluation scheme reflects the school or class population. Good practice dictates that courses be offered at different levels of difficulty to accommodate varying levels of ability. By its very nature, the kind of evaluation described here is student-oriented, in that each student knows the criteria and other necessary information (such as the audience) before writing a word. Ideally, the scheme will produce a set of standards that are clear to everyone.

This scheme will adapt itself to a wide range of reporting techniques within a school or system; marks, letters, anecdotal reports, or, ideally, a combination of these may be used. The scheme also presupposes full staff involvement in the development of writing skills, and must be flexible enough to have validity across the curriculum.

A Co-operative Scheme

The evaluation of student writing should be a co-operative effort. A student must accept responsibility for his or her acquisition of writing skills. By ensuring that the ongoing evaluation of student writing is done by the student, his or her peers, and the teacher, and by

insisting that each of these has a significant bearing on day-to-day classroom procedures, we develop the atmosphere of a writing workshop. In this atmosphere, the acquisition of writing skills becomes a co-operative activity.

Similarly, when summative evaluation is necessary, the kind of reliability produced by all these evaluators, in combination with writing samples, criteria-mastery charts from the writing folder, and some random sampling or full grading through holistic, multiple marking, goes far beyond what frequent teacher grading alone can produce. The often-criticized unreliability of the grading of writing should then disappear.

A Prismatic Scheme

An effective evaluation scheme reflects in a variety of directions. It reveals student achievement, progress, strengths, and weaknesses. It reveals teacher methodology, perception, and effectiveness. Program suitability, range, and impact should also be apparent in the evaluation pattern.

Mode, Audience, Purpose, Register

We write for a variety of audiences and purposes and in a variety of modes. The young writer decides on the register of language demanded by these factors. Any set of criteria established for the evaluation of writing must then include some consideration of how well the writer achieved his or her goals in terms of audience, mode, register, and purpose. By focussing on these facets, we emphasize the degree to which the student was successful in using language appropriately and effectively. In this way, student motivation and confidence are encouraged.

The Student Writing Folder

The student folder is a natural outcome of an evaluation system aimed principally at formative evaluation. It is more than a collection of the student's work over the year; it is an active teaching and evaluative tool, which teacher, student, and peers use frequently as part of the process of improving writing. Along with all the student's writing, it should contain clearly stated criteria on which the ongoing monitoring process may be based; a spelling list (preferably student-produced) in which problem words are explored; some kind of usage-mastery chart or graph, by which specific progress in eliminating problems is demonstrated; and comments by the student, the student's peers, and the teacher about writing progress in these areas.

When summative evaluation is done, the teacher and student may select from the folder a sample of work for grading. As well, at the end of the term or year, a portfolio of student samples, comments, and mastery information should ideally accompany the student to the next teacher of writing.

Specific Criteria

If there is one factor central to the informed evaluation of student writing, it is the development of specific criteria. The advantage of an evaluation scheme based on criteria that are both known and understood by teacher and students (and the rest of the educational spectrum) far outweighs any potential disadvantage.

Many sets of criteria have been advanced; more appear every day. The Intermediate Division English guideline includes an ability-scale rating approach. This checklist illustrates the flexibility of the criteria in that it allows additional factors to be added as the student masters certain areas.

The guideline also presents an analytic scale, based on wide research, which includes five basic criteria for evaluating writing – ideas, organization, diction, style, and usage; and another rating scale using a range of 1 to 5. Additionally, two separate criteria are elaborated on pages 112-13 of the guideline.

A practical application of some of these ideas can be found in C. R. Cooper, "Tonawanda Middle School's New Writing Program", pages 56-61. This article contains a further set of criteria and writing strategies.

Holistic, Multiple Marking

The employment of multiple markers using a holistic approach is certainly not a new idea.[2] As a method of summative evaluation, it has proven to be remarkably accurate. Used sparingly and intelligently, it can provide a measure of student growth in writing that is readily acceptable and meaningful for students, parents, administrators, trustees, and the public. More significantly, it can reduce the need for frequent summative marking procedures and allow teachers more time to teach writing skills on an individual basis.

The process is described in the Intermediate Division English guideline, pages 113-14, and further details appear in Appendix B, pages 126-27 of the guideline. It is also described in considerable detail by Charles Cooper in *Evaluating Writing: Describing, Measuring, Judging*.

The principal advantage of holistic, multiple marking is that it removes the teacher from the role of sole evaluator and replaces him or her with a small team of teachers, who evaluate anonymous papers, and are, in turn, unknown to the student. This method allows the teacher and the student to function as a team, whose goal is to produce a noticeable improvement in their holistic score. If there is no noticeable growth in student writing skill through the term or year, then very serious attention must be paid to the nature of the writing program.

The participation in holistic scoring of teachers from across the curriculum (in the secondary panel particularly) is a good way of involving all staff in the development of writing standards throughout a school.

The following experiment may be considered during a secondary school holistic marking session: Include writing samples from the whole school population. The work should not be separated into basic, general, advanced, or enriched categories. When all writers are anonymous, does anything happen to their scores? Does any measurable progress in writing achievement show up in the next session?

[2]J.N. Britton, N.C. Martin, and H. Rosen, *Multiple Marking of English Compositions*, Schools Council Examination Bulletin no. 12 (London: Her Majesty's Stationery Office, 1966).

Diagnostic Procedures

There is a need for more sophisticated procedures for the assessment of student skill levels at the beginning of a term than are presently in general use. The writing folder, if fully adopted, would solve many of the problems faced by all English teachers in September. The folder could present to the new teacher a relatively clear picture of the student, especially if the student and his or her former teachers have conscientiously maintained the cumulative records.

If the carrying forward of a detailed writing folder is not workable, then other diagnostic steps should be taken. Generally, initial assessment has involved a piece of writing, and generally the teacher has thus obtained a picture of the overall level of the writer being assessed. It is suggested here that the teacher make that early writing a source of information about more than just general ability levels. Some close analysis of some of these early pieces of student writing should be done.

There are very complicated procedures for assessing various aspects of writing. Several arrive at a conclusion about syntactic maturity through an analysis of clause use, sentence length, types of clauses used, use of gerundial or infinitive phrases, errors per 300 words, and so on. These procedures can be simplified and adapted to a diagnostic procedure involving the specific criteria to be pursued throughout the course.

For example, Intermediate Division teachers or an English department in a school could assess the student population in English early in the year by noting some of the following factors, above and beyond the holistic impression obtained:

- the number of clauses used for every ten sentences;
- the average length in words of clauses;
- the number of gerundial or infinitive phrases used;
- the number of sentences per hundred words;
- the total error count;
- the number of spelling errors per hundred words;
- the kind of errors encountered.

This information could be shown on an error chart.

Such analysis, while time-consuming, would give all teachers a clearer view of some of the areas needing work – and, perhaps more importantly, some of the areas not needing work.

For more detailed information about this kind of analysis, see C. R. Cooper, "Measuring Growth in Writing," pages 111-20; F. O'Hare, *Sentence Combining: Improving Student Writing Without Formal Grammar Instruction,* pages 5-18; and K. W. Hunt, *Grammatical Structures Written at Three Grade Levels.*

Writing Workshop

There are many ways to organize a writing workshop approach. However, the critical aspect of organization lies in the attitudes which a particular classroom organization stimulates. Students who see writing as a process encouraged, stimulated, and monitored through their own participation and that of their peers and their teacher are more likely to have a positive attitude towards writing.

The writing workshop responds to a variety of known factors about the acquisition of skills:

- It allows students to talk about projects, and to manipulate syntactical patterns in questioning, formulating, rephrasing, and responding.
- It allows students to test responses to preliminary composing ideas within the group.
- It affords the young writer, who eventually must write alone, the opportunity to take to that solitary task a range of ideas, reactions, and other stimuli that he or she otherwise would not have.
- It allows the teacher an opportunity to evaluate a great number of the factors involved in the writing process, particularly the prewriting and postwriting activities that occur within the boundaries of the writing workshop.

Bibliography

Britton, J. N.; Martin, N. C.; and Rosen, H. *Multiple Marking of English Compositions.* Schools Council Examination Bulletin no. 12. London: Her Majesty's Stationery Office, 1966.

Cooper, C. R. "Measuring Growth in Writing". *English Journal* 64 (March 1975): 111-20.

_____. "Tonawanda Middle School's New Writing Program". *English Journal* 65 (November 1976): 56-61.

Cooper, Charles, and Odell, Lee, eds. *Evaluating Writing: Describing, Measuring, Judging.* Urbana, Ill.: National Council of Teachers of English, 1977.

Dixon, John. *Grow Through English.* London: Oxford University Press, 1967.

Evans, P.; Brown, P.; and Marsh, M.; eds. *Criteria for the Evaluation of Student Writing, Grades 7 and 8: A Handbook.* Toronto: Ontario Institute for Studies in Education, 1976.

Hunt, K. W. *Grammatical Structures Written at Three Grade Levels.* National Council of Teachers of English Research Report no. 3. Champaign, Ill.: NCTE, 1963.

Kirschenbaum, H.; Napier, R.; and Simon, S. B. *Wad-ja-get?: The Grading Game in American Education.* New York: Hart Publishing, 1971.

O'Hare, F. *Sentence Combining: Improving Student Writing Without Formal Grammar Instruction.* National Council of Teachers of English Research Report no. 15. Champaign, Ill.: NCTE, 1973.

Ontario Ministry of Education. *Evaluation of Student Achievement: A Resource Guide for Teachers.* Toronto: Ontario Ministry of Education, 1976.

Writing Mechanics – 1969-1974. National Assessment of Educational Progress. Washington: U.S. Government Printing Office, 1975.

Writing: National Results – Writing Mechanics. National Assessment of Education Progress, Report no. 8. Washington: U.S. Government Printing Office, 1972.

Language Study and Evaluation

Introduction

Language objectives recur throughout the Intermediate Division English guideline, and the development of the student's proficiency in language is stressed throughout that document. The chief purpose for including a separate section here is to emphasize the importance of language study by illustrating means by which the measurement of language development and language appreciation can be achieved.

The student's understanding and appreciation of language should not be restricted to the standard dialect; indeed, emphasis must be placed on the development of the student's appreciation of dialects other than his or her own and on the student's understanding of the process of language change. The student should be aware of the effects of language, not only of malapropisms and inappropriate deviations from standard usage, but also of the use of language in influencing thought.

Explorations in diction should stress the subtlety of language use – the nuances of different word choices – as well as precision and colour. Above all, emphasis should be placed on the relationship of both oral and written language to purpose and audience. The focus should be on the language the student generates – language in use.

It is therefore difficult to separate the measurement of achievement of language objectives from the total English program. Fundamentally, it matters very little whether the student can cite seventeen uses of the comma or can recognize comma blunders in exercise material. What does matter is whether the student makes these errors in the written work to which he or she has a commitment.

Language Fundamentals and Writing: The Individual

Most of what needs to be said about student language performance in written English has already been outlined in the section on writing. Measurement of language performance generally should only be undertaken on the basis of a wide sampling of the student's production of both oral and written language.

The activity of student labelling of errors in exercises composed of badly written sentences divorced from any context – the type of exercise found in altogether too many textbooks – has little to commend it. A fundamental principle of sound language teaching may be undermined by concentration on such exercises; the student should be examining models of excellence, not concentrating on English of poor quality.

As already suggested, the writing folder is an effective means of observing the student's strengths, weaknesses, and development as a writer. In the context of that folder, recommendations are also made concerning the use of criteria, including those relating to spelling, punctuation, and subject-verb agreement. Further suggestions are made concerning checklists or mastery charts for a number of readily checked areas of language mastery. Such checklists may be maintained to a large extent by the student, not only to reduce the burden of marking for the teacher, but also to encourage the student's positive involvement in assessing and improving his or her skills.

Implied here is the recommendation that the student's "correction" of language should be centred largely on personal writing, or, in writing groups, on the writing of peers. In short, the student should be focussing on those personal writing problems that require attention instead of reviewing the multitude of possible errors that he or she may make.

Emphasis has been placed on positive criticism and encouragement in the marking of student work. As well, suggestions have been made in the Intermediate English guideline concerning improvement in the dialogue between teacher and student by close agreement concerning writing modes and expectations and by limited use of abstract or esoteric symbols.

It must be stressed that students at any grade level will reflect a very wide range of competence in language. Teacher expectations must be realistic. Furthermore, it is not helpful to the weaker student to have all his or her linguistic sins exposed at once. To "red circle" every error, major and minor, may defeat the whole intent of the interchange between teacher and student. Teachers must be selective, concentrating on the most fundamental problems first. They should concentrate as well on reading the writing which the student has edited and revised.

Achievement in most of those objectives concerned with increased mastery – of such conventions as spelling, punctuation, and capitalization as well as of usage, grammar, and acceptable syntax – can best be measured through information supplied by the writing folder. The folder is also the best practical source of diagnostic information concerning individual language needs.

In addition, it is principally through the student's writing that improved performance in language choice and flexibility of style in relation to purpose and audience can be assessed. Some practical suggestions concerning the use of such things as criteria and a simple analytic scale have already been offered. Here, three cautions regarding the use of criteria deserve repetition:

1. Limit the criteria for any assignment to only a few.
2. Be certain that the criteria employed are both understood by the student and appropriate to the mode of writing or the demands of the assignment.
3. In selecting criteria, maintain a balance between evaluating the mechanical aspects of writing (spelling, grammar) and evaluating the positive features of style and diction.

Language Fundamentals and Writing: The Group

So far, techniques for measuring student performance in language have focussed on the individual and particularly on diagnostic and formative evaluation. Much valuable information can also be gained concerning group performance and overall patterns of achievement in language through occasional, careful sampling of student writing at a particular grade level within a class, a school, or a system. Such information, systematically gathered, provides useful data both for planning language programs and for assessing their overall effectiveness.

Holistic, multiple marking is one method of obtaining this information. (See the discussion of holistic scoring in this document, in *English, Intermediate Division, 1977,* and in Paul Diederich, *Measuring Growth in English,* pages 113-14.) Holistic scoring provides an opportunity for obtaining a large and valid sample of student writing, though, of course, a sample can be taken from regular assignments at any point in the year.

There are many measures – some straightforward, some rather complex – that can be made of general student competence. The error count is, perhaps, the most obvious. Even here, though, a training session is required for scorers to come to reasonable agreement about classification of errors; in addition, at least two scorers should count the errors in each composition. An account of a technique for making error counts may be found in R. Traub et al., *Secondary-Postsecondary Interface Project II: Nature of Students.* A somewhat similar technique – for use with the writing of nine-year-olds, thirteen-year-olds, and seventeen-year-olds – is described in the reports on the evaluation of writing of the National Assessment of Educational Progress for 1969 and for 1974 (see the bibliography to this section). Both of these reports also provide some information about the likely frequency of certain types of errors.

This information is likely to be generally helpful both in planning a program and as a means of estimating the success of the program. Teachers may find, however, that some general classifications of errors need to be refined further. There is an obvious difference in language control between students who write "They was too small" and those who write "He was one of the boys who was too small," though both errors can be classified under "subject-verb agreement"; there is a similar distinction to be made between students who write "beleive" and those who write "pyschology", though both have made spelling errors.

In this same context, it should be noted that as students improve in their ability, or at least their willingness, to employ complex syntax or sophisticated vocabulary, the anticipated drop in the absolute frequency of certain classes of errors may not occur, though in fact language development may have been considerable.

In short, evaluators should proceed carefully. The error count, used with care and attention, can provide a good deal of useful information related to the development of the language program.

Measures of syntactic complexity (when applied to a group and for the same writing mode) are also an important index of development. They may be as straightforward as observation of mean sentence length or they may measure frequency of use of phrases, introductory clauses, or verbs, or variety in sentence order. A number of strategies are described in the following publications: the report of the National Assessment of Education Progress (referred to above); Walter Loban, *Language Development: Kindergarten Through Grade 12*; C. R. Cooper, "Measuring Growth in Writing", pages 111-20; and C. R. Cooper and Lee Odell, eds., *Evaluating Writing: Describing, Measuring, Judging.*

The utility of measures of syntactic fluency depends on the accumulation of information over time. A student who measures significantly below the group – unless that student is experimenting in Hemingway's style while the rest are imitating Milton or Henry James – is likely to be in trouble. However, the main value of such measures is to gauge net gain over time, and, used in this way, they are a convenient measure of program effectiveness.

Specific Testing of Language Competence

Cautions have been expressed concerning language tests where sentences are divorced from context and where, too frequently, the emphasis is on sentences that are error-filled rather than on models of effective structure.

There may be occasions where the group's or, more particularly, the individual's grasp of a specific feature of grammar, syntax, or style must be tested. How can the teacher best test this competence, avoiding a negative approach and at the same time relating performance on the test as closely as possible to the student's actual everyday writing performance? It is certainly true that a good performance on a language test provides no guarantee of carry-over to actual writing performance. In this regard, *the best course to follow is to develop test items that compel the student either to construct a sentence or to choose the best from a number of reasonably acceptable alternatives.* The emphasis should be placed on synthesis and on the most logical or most effective thought sequence.

In the following examples of test items, the principle is to have students demonstrate their understanding by *constructing* their own sentences.

Connectives

- Use the word "nevertheless" between two thoughts. Punctuate correctly.
- Write two complete thoughts, the second of which begins with the word "subsequently" or "finally".
- Write a three- or four-sentence paragraph that uses such sentence adverbs as "therefore", "moreover", or "subsequently".

Parallelism

- Write a sentence using "neither" and "nor", with at least three words in parallel order following each.

Punctuation

- Write two sentences. In one, use the semi-colon correctly and in the other use the colon correctly.

Spelling of Homophones

- Use "they're", "there", and "their" in a single sentence.
- Use "its" and "it's" in a single sentence.

Sentence Structure

- Complete the following sentences, which do not yet express a complete thought.

a) All day a cold wind from the north . . .
b) The watchers anxiously looked at . . .
c) Along the eastern horizon great thunder clouds . . .

Although students are working in areas in which they may be having difficulties, in these examples their

attention is fixed not on identifying errors, but on constructing an effective sentence. A small bank of tests like the following might be developed as a follow-up for students who are having problems writing proper sentences:

Fused-Sentence Problems

– Create single sentences in each of the examples below, inserting where appropriate the proper punctuation and/or connective, and capitalizing where necessary.

a) She was very angry with me.
I was angry with her too.

b) The boys sat very still.
They did not want to be heard.

c) My Irish setter has beautiful red hair.
Your shaggy mutt is a disgrace.

d) Never have I seen such a mess.
Peanut butter and jam were all over the table, the chairs, and the floor.

The specific problem (connectives, or the tendency of students to ram sentences together without punctuation) is emphasized here. The following procedure is recommended for dealing with such specific writing problems:

– Observe the type of error that the particular student frequently makes.
– Administer a quick test (as above) to see whether the student understands the principle involved.
– If the student does not understand the principle involved,
  a) provide remedial attention;
  b) administer a quick retest (as above) to see whether the principle is grasped;
  c) carefully check (and have the student check) further writing assignments for a recurrence of the problem.

Additionally, exercises in sentence-combining can be an effective way of improving students' sentence construction. A number of lessons and problems in this area are outlined in F. O'Hare, *Sentence Combining: Improving Student Writing Without Formal Grammar Instruction*. Using these as a model, teachers could construct specific exercises for sentence-building by drawing from effective sentence construction in the literature their students are studying or ineffective constructions in the students' own writing.

A further rich resource for exercises in language and writing is Littleton Long, ed., *Writing Exercises from "Exercise Exchange"*.

The task of developing short tests and exercises for many specific language needs is quite simple and straightforward for the individual teacher. However, exercises and tests that treat complex problems in a systematic way (as O'Hare does with exercises in connecting sentences) require much time, informal research, and patience. It is therefore advantageous for all staff, across several grade levels, to combine efforts in producing tests and exercises that respond to typical problem areas and to keep an active file of such diagnostic and mastery tests, which can be reused from class to class and year to year.

Explorations in Language

The exercise materials illustrated thus far have focussed chiefly on sentence construction, spelling, and punctuation. According to the objectives identified at the outset of this section, students are expected to develop an awareness of language development, registers appropriate to different contexts, and dialect variation. Diction and imagery are also to be studied, for example, in the context of the effect or power of language in influencing thought.

It is difficult to outline specific evaluation strategies for language (apart from those alluded to in the criteria for the evaluation of writing itself) without at the same time delving in great detail into language principles, resources, and teaching strategies.

One of the most imaginative collections of strategies for explorations in language is G. A. Tilly, *Connutics: A K-13 Language Resource Book*. Also to be especially commended are the exercises and activities in the Language of Man series, published by McDougal, Littel, and in the Language Matters series, published by Thomas Nelson.

For material on the derivation of words and the connotation of language, see especially J. F. Littel, ed., *Gaining Sensitivity to Words*. As well, J. F. Littel, ed., *How Words Change Our Lives*, deals with the impact of language, using the media and propaganda as examples.

All these resources have in common an exploratory, empirical approach to language (descriptive rather than prescriptive). In general, appropriate evaluation strategies will emphasize activities or projects with language, in the form of reports, creation of dialogue in a consistently appropriate register, the writing of effective advertising copy, and so on. The study of language for students in the Intermediate Division should lead directly to activities that illustrate whether the students are able to put to use what they have learned. The resulting student work can then be evaluated by the same criteria that are used with other forms of writing.

The matter of language change and those influences (such as other languages, varieties of English such as British, American, or Canadian, neologisms from science and technology, the jargons of sports) that bring about change are subjects that most young people find very interesting. A particularly valuable resource in this regard is the survey of Canadian English undertaken in 1971-72 by the Canadian Linguistics Association and the Canadian Council of Teachers of English. The usages surveyed and the pattern of results according to age (parent or student), sex, and province are reported in *The English Quarterly* 5 (Fall 1972), and have been published separately as well by the Canadian Council of Teachers of English. Another useful resource is the text *Our Own Voice: Canadian English and How It Is Studied* by R. E. McConnell.

The patterns of results on numerous items in the 1971 survey provide interesting food for discussion about language. A valuable project for a class or class group would be to design its own language survey and to analyse the results. Student explication of the results of this survey, as well as the case made for the inclusion of particular items, is the principal means by which the teacher can determine student understanding of

registers, dialect, language change, and sources of influence on language development.

Here are three examples from the survey of Canadian English, with total responses (shown as percentages):

Item: (A) Everybody gets their reward.
(B) Everybody gets his/her reward.
(C) Either one

	A	B	C
Male Parents	55	29	15
Female Parents	50	34	16
Male Students	61	18	21
Female Students	60	18	22

Item: (A) If he was here, things would improve.
(B) If he were here, things would improve.
(C) Either one

	A	B	C
Male Parents	38	50	11
Female Parents	30	61	9
Male Students	44	33	22
Female Students	42	40	18

Item: The first syllable of *lieutenant* is pronounced like (A) left; (B) loot; (C) either way.

	A	B	C
Male Parents	43	48	8
Female Parents	35	56	9
Male Students	17	71	11
Female Students	11	78	10

Any of these examples, and many more, are suitable for use with students. They can provoke fascinating discussion about language and change, and they can suggest further, similar items.

Expansion of Vocabulary

Writing opportunities and exercises that focus on precise or effective connotative language are undoubtedly the best means of encouraging students to make full use of the resources of the language and of appropriate aids to its use – models of writing, the dictionary, the thesaurus.

Unless some topic, lesson sequence, or particular reading selection demands the understanding of new vocabulary, most testing of vocabulary "expansion" is rather arbitrary and fragmented and does not usually involve words that students actively use. In vocabulary testing in standardized tests, words are often taken out of context and nuances of meaning ignored. There are exceptions, however: The *Co-operative English Test* (Princeton, N.J.: Education Testing Service, 1950) presents some interesting items (which may provide teachers with ideas for the creation of "short test" items) that test the student's effective control and precision in the use of their current vocabulary. Test security forbids reproduction of some items, but below are a few, modelled on the *C.E.T.* Where a case can be made for more than one answer, depending on tone and context, much refinement of language learning will take place in the follow-up discussion about preferences.

1. Objective: connotation and precision
Many cities, though harbouring slums or deteriorating housing in the core, are ( ) for their redevelopment of the business district.
a) noted
b) admired
c) notorious
d) appreciated

2. Objective: connotation
Leaning against his shoulder in the moonlight, she ( ) the warmth of her feeling for him.
a) whispered
b) murmured
c) babbled
d) expressed

3. Objective: cliché and register
( ), she is one of the finest young ladies I have met.
a) In the final analysis
b) When the chips are down
c) When all is said and done
d) In sum

4. Objectives: (i) register, cliché, euphemism; (ii) precision and colour
Shortly after my uncle ((i)), his relatives began ((ii)) over his small estate.

(i)	(ii)
a) passed away	a) debating
b) entered into rest	b) arguing
c) kicked the bucket	c) squabbling
d) died	d) disagreeing

Once again, the development of short, easy-to-score tests of language should be an activity shared by the staff as they develop units in such areas as language change, register (formal-colloquial), diction, and imagery. The tests may be used directly as measures of achievement of specific objectives, as reinforcement, and as teaching materials in their own right.

Oral Language

Most of this section has focussed on written English, both because of the current preoccupation with student performance in writing and because, in writing, students provide us with the most readily measurable example of their language performance.

Written English, in its various registers, is different in numerous respects from spoken English. It is important that students (and teachers) be aware of this distinction. Much of what has already been said about language development, influences on language change, and register is pertinent to spoken language. Differences between written and oral English can best be learned in the context of activities in which comparisons are made – activities such as those in which a story is recreated in dialogue, a written advertisement is compared to its television or radio version, or a letter of application is compared to an interview. Negative comments, such as "That's wrong usage" or "That's unacceptable English", in response to students' spoken language should be avoided.

Many activities involving the study of written and spoken English may be found in the publication *Connutics,* referred to earlier.

Most of the objectives for evaluation set out at the beginning of this section have equal pertinence to spoken and written English. Unfortunately, oral language – the principal way in which language is used – is often neglected in teaching English and in the evaluation of language performance, because the record of performance is so impermanent. Likely the best way to underline the importance of performance in spoken English is to ensure that the evaluation of speaking is included in the evaluation of oral language activities such as group discussions, individual presentations, and dramatic improvisation. Records can be kept of students' participation in discussion, as well as of their performance in more formal presentations such as speeches, oral reports (formal and informal), and debates.

Though one continually faces the problem of the transience of oral work, classroom presentations are an excellent opportunity for peer evaluation using numerous criteria (see the Intermediate Division English guideline, page 118, for a simple chart). With the help of the tape or cassette recorder, self-evaluation can also take place.

Bibliography

The following materials are especially recommended:

Barnes, D. *Language, The Learner, and the School.* Harmondsworth, England: Penguin Books, 1969.

Braddock, Richard Reed; Lloyd-Jones, Richard; and Schoer, Lowell. *Research in Written Composition.* Champaign, Ill.: National Council of Teachers of English, 1963.

Cooper, C. R. "Measuring Growth in Writing". *English Journal* 64 (March 1975): 111-20.

Cooper, C. R., and Odell, Lee, eds. *Evaluating Writing: Describing, Measuring, Judging.* Champaign, Ill.: National Council of Teachers of English, 1977.

Diederich, Paul. *Measuring Growth in English.* Champaign, Ill.: National Council of Teachers of English, 1974.

Doughty, P., and Thornton, G. *Language Study, the Teacher and the Learner.* London: Arnold, 1973.

Halliday, M. A. K. *Explorations in the Function of Language*. London: Arnold, 1973.

Hunt, K. W. *Grammatical Structures Written at Three Grade Levels*. National Council of Teachers of English Research Report no. 3. Champaign, Ill.: NCTE, 1965.

Littel, J. F. *How Words Change Our Lives*. Evanston, Ill.: McDougal, Littel, 1971.

Littel, J. F., ed. *Gaining Sensitivity to Words*. Evanston, Ill.: McDougal, Littel, 1973.

Loban, W. D. *The Language of Elementary School Children*. National Council of Teachers of English Research Report no. 1. Champaign, Ill.: NCTE, 1963.

Long, Littleton, ed. *Writing Exercises from "Exercise Exchange"*. Champaign, Ill.: National Council of Teachers of English, 1976.

McConnell, R. E. *Our Own Voice: Canadian English and How It Is Studied*. Toronto: Gage Educational Publishing, 1979.

O'Hare, F. *Sentence Combining: Improving Student Writing Without Formal Grammar Instruction*. National Council of Teachers of English Research Report no. 15. Champaign, Ill.: NCTE, 1973.

Tilly, G. A. *Connutics: A K-13 Language Resource Book*. North York: Board of Education for the Borough of North York, 1972.

Traub, Ross; Wolfe, Richard; Wolfe, Carla; Evans, Peter; and Russell, H. H. *Secondary-Postsecondary Interface Project II: Nature of Students*. 2 vols. Toronto: Ontario Ministries of Education and of Colleges and Universities, 1976.

Wilkinson, A. M. "The State of Language". *Educational Review* 22 (November 1969). Available from the NCTE.

Wilkinson, A. M., ed. "The Context of Language". *Educational Review* 23 (June 1971). Available from the NCTE.

Writing Mechanics – 1969-1974. National Assessment of Educational Progress. Washington: U.S. Government Printing Office, 1975.

Writing: National Results – Writing Mechanics. National Assessment of Educational Progress, Report no. 8. Washington: U.S. Government Printing Office, 1972.

Evaluation: School, Program, System

Introduction

The previous sections identified and illustrated a wide variety of evaluation strategies, most of them of use to the teacher right in the classroom. In this section we are attempting to look beyond the classroom to the school and to the school system. Assuming that there is a wide interest in the evaluation of English and that school systems are being asked to provide information about performance in English, how can a good selection of strategies related to important program objectives be systematically assembled to provide this information?

Essentials for a Systematic Approach

The following are the essential underpinnings of any systematic approach to the evaluation of English:

1. *An approach to evaluation in English is valid only if it is referenced to the principal objectives of the English program.* (Suitable objectives are listed on pages 10-13 of the Intermediate Division English guideline.)
This principle has several implications:
a) At the initial stage of course development, teachers must be involved in establishing what the principal objectives of the English program are to be.
b) These objectives must include affective as well as cognitive aspects, oral as well as written English, and receptive skills (reading, viewing, listening) as well as expressive skills (speaking, writing).
c) Teachers must make decisions about appropriate measurement strategies for each objective. Some of these strategies will be more straightforward than others, but all of them must find their place in a balanced evaluation of student performance.
d) Opportunities should exist for the sharing of strategies among teachers, among schools, and among systems. This may be done through the joint development of tests and pools of test items, through the systematic, regular "holistic" scoring of samples of student writing across several grades, or through an exchange of strategies such as checklists and charts (and training in their use).

2. Since the development of an effective approach to evaluation embraces so many objectives and strategies, *any acceptable approach to evaluating English should have its focus first of all within the school* (or family of schools). Teachers should have the primary say in what should be evaluated and they should work together in deciding which strategies are likely to be most effective. With whatever support is possible from principals, consultants, and school administration, teachers should be responsible for the establishment of a reasonably systematic approach to evaluating student achievement and the program.

An Evaluation Program Within the School

An evaluation program should have the following characteristics:

1. It should provide a balanced, effective plan for the measurement of student achievement, with reference to major objectives of the program. This embraces diagnostic, formative, and summative evaluation.

2. It should supply measures that will assist teachers and consultants in assessing program effectiveness and in making decisions with respect to changing methodology, content, or balance.

3. It should contain a "routine" or checklist for use by principals, teachers, or superintendents for ensuring that resources are available and that classroom procedures are sound for the conduct of the program.

Each of these requires, as a first step, reasonable specification of the objectives of the English program and of the expectations for particular grade or program levels. The planning of revisions to current courses of study to bring them into conformity with the Intermediate Division guideline or decisions concerning modifications are dependent on clear agreements about objectives.

This basic planning will ensure that a balance is achieved within the program among listening, speaking, reading, writing, and literature activities, and that opportunities for integrating a number of language activities within different units of the program are effectively exploited.

Some of this general planning is appropriate at the system level. Within the school, however, specific and detailed decisions are required concerning content, resources, and appropriate evaluation strategies, with, very likely, considerable variety from school to school and level to level. In the context of this second stage of program planning, some common decisions should be reached about such details as the use of writing folders and checklists, the keeping of records, and the employment of reporting strategies that reflect the variety of objectives and attainments. Here also recognition must be given to the fact that Intermediate Division students are characterized by a wide range of achievement levels in the various aspects of language development.

As tentative decisions are reached about types of evaluation strategies, their purposes, and the frequency of evaluation, teachers should begin to build short tests (where these seem appropriate) to determine achievement in specific objectives, particularly in the areas of reading and language study. These tests should each be limited to a single, quite specific aspect of performance, and, as they are used, acceptable levels of achievement can be defined.

Out of such short tests, more complex, combined tests can be constructed for general use within a system, and item pools can be developed. The emphasis here should be on shared test development and on the primacy of teachers in determining what should be tested. Although the refinement of any test or testing strategy does call for the assistance of people trained in test design and measurement, the initial design, as well as recommendations for use or modification, should remain in the hands of the subject specialist.

As strategies are identified, developed, and modified, with teachers trained in their consistent use, the evaluative process itself will provide considerable information about reasonable student performance levels, so that test strategies will increase in value as diagnostic instruments for individual students and for class groups. Also, as test (or other strategy) performance information is accumulated, the tests will become increasingly useful as measures of program effectiveness and as a source of information for program change.

A Checklist for Teachers, Principals, Consultants, and Administrators

There are a number of questions principals and administrators should ask when assessing the effectiveness of a school program in English or when observing an English classroom.

The following questions may serve as suggestions, by no means exhaustive, for the continuing in-school assessment of courses, course development, and program. Using some of these as a basis, teachers, principals, and department heads might compose questions of their own to aid them in observation and in regular program review.

1. The Intermediate Division English guideline states that "the *basis of the compulsory English curriculum is an integrated balance of reading, writing, speaking, and listening,* and it is upon this foundation that courses should be based" (page 4). Do courses at each grade level reflect this "integrated balance"? For example:
- Are such topics as "grammar", "spelling", "language", and "writing" treated as separate exercises or are they tied closely to the student's writing performance?
- Are language problems or "basic writing skills" treated as separate exercises or are they tied closely to the student's writing performance?
- Is the timetable fragmented, with different teachers responsible for different components of English?
- Is reasonable balance maintained between literature and the language program?

2. Have courses been designed with specific reference to the objectives outlined in the Intermediate Division English guideline (pages 10-13)? For example:
- Do course objectives at each grade level relate specifically to each of the areas defined in the objectives outlined?
- Is there some sequencing and balancing of objectives (and, hence, of course content) from one grade level to the next? To what extent is this sequencing and balancing based on evidence of student language achievement and needs at different levels?
- In course outlines, has reference been made to specific evaluation strategies or instruments to be employed? Are these, in turn, referenced to specific objectives with some indication of appropriate criteria for performance?
- Are recording and reporting strategies compatible with these? Is it possible to transmit reasonably specific information to the student, the parent, and the system?

3. Is there a schedule for course review and revision? For example:
- Is there a reasonably firm information base upon which to make recommendations for revision and to monitor the effects of the revision?
- Is there a teacher chiefly responsible for course co-ordination and review for each grade or program level?

– Are teachers of particular courses encouraged to meet occasionally to discuss course needs, problems, and resources?

4. a) Does each teacher have on hand (or readily available to him or her):
– the pertinent Ministry of Education guideline;
– all Ministry support documents;
– a curriculum guide and/or support materials (where these are provided within a system);
– means for gaining access to such resources as films, videotapes, the school's or board's resource centre;
– several copies of a variety of current language texts approved in *Circular 14* for use with different groups of students;
– a small library of current books and resources (such as those produced by the National Council of Teachers of English) for the teaching of English?

b) To what extent are these resources used and found to be useful?

5. What kinds of activities take place within the classroom? For example:
– Is there provision for students to work frequently in groups and on an individual basis?
– Is some free reading time available and are attractive reading resources at hand or close to hand?
– Does oral language receive attention?
– Do studies in literature demand from the student only recall of contents, identification of literary terms, and literary criticism, or do they call, through a variety of activities, for student response to literature?
– Is the student encouraged to maintain records of his or her participation, reading, and writing, and to assess his or her progress?
– Does the study of language (usage, grammar, vocabulary, register, etc.) take the form of drills and exercises, or is it integrated with language and, especially, writing performance?

6. What approaches are taken by teachers towards evaluation? For example:
– Do teachers participate in the selection of all testing instruments brought into the classroom? Are they informed about the limitations (both in accuracy for reporting individual scores and in objectives tested) of the tests used?
– Do teachers use a variety of strategies to take into account student attitude as well as knowledge?
– Is there opportunity for teachers to share in the development of test instruments and strategies both in order to encourage greater consistency in their use and to build up a variety of instruments that can be exchanged?
– Do teachers use and, from grade to grade, take advantage of the cumulative writing folder and the information it contains?
– Is opportunity provided for the accumulation of information on student performance (tests, checklists, attitude scales, writing sampling, etc.) to improve the quality of the instruments and methods used in evaluation and to serve as a data base for making judgements about student performance in English? Can this information be cycled back into course and program modifications?

Testing and Sampling in a School System

The selection and design of evaluation instruments and strategies should begin in the school. The strategies used should be *varied* in terms of type and of evaluator (the student, the student's peers, or the teacher) – and should reasonably reflect important program *objectives*. From this basis, it is possible to move constructively towards evaluating student achievement across a program, a system, or the province. Such testing programs should be low-key and not perceived as a threat to the teacher or the program. If an evaluation plan is perceived as being imposed on the teacher or at cross-purposes to the teacher's understanding of what is important in the program, it is unlikely to be effective. Fortunately, there are relatively easy ways of avoiding a testing program that may be a traumatic experience for all. For example, there is no need to test every pupil in every class on every objective. Student sampling and matrix sampling are less expensive and less threatening alternatives where the purpose is to *monitor* program and student achievement.

A reliable and valid means for periodic measurement of writing achievement (within a grade, across grades, or across a system) has already been outlined in some detail earlier in this document. If this approach is undertaken systematically, a great deal of sound information will be at hand concerning the state of student writing competence at different grade levels. This information can also be used to measure gains resulting from program modification.

There are so many possible objectives for an English program that means ought to be found for developing in a reasonably systematic way a variety of measures of performance. Most current standardized group tests of English provide a rather poor fit with curriculum objectives, and frequently the testing strategies are in conflict with what are seen as appropriate approaches to language learning. A useful alternative is the development of tests by practising teachers, working co-operatively.

Criterion-Referenced Testing

The crucial difference between a norm-referenced test (which most standardized tests are) and a criterion-referenced test is that the former is used to compare the performances of students, while the latter is used to measure the individual student's success in meeting a specific objective.

While there are occasions when it is desirable and legitimate to compare performances (as in tests predictive of success in a future course or program), for many evaluative purposes norm-referenced tests are inappropriate.

> One cause for the inappropriateness of norm-referenced measures is that these instruments must produce variant scores. If variability is not present in the responses of those who take the test, then those responses cannot be contrasted. And remember that the purpose of norm-referenced tests is to permit comparisons among people. Thus, a major thrust of norm-referenced test construction and test revision is to produce variability. Instructional specialists are frequently working in the opposite direction. The good teacher would often like to

> reduce variability by getting all of his pupils to display a given level of excellence. Surely there will be differences among learners, but with respect to fundamental objectives in reading, we can strive for 100 per cent mastery by all pupils. To put it another way, the educational evaluator should be interested in how many learners can achieve an educational objective, not how the learners compare with each other.[1]

Dr. Popham goes on to describe how items on norm-referenced tests are manipulated to produce increased variability or score spread, often at the expense of the concept being measured. Overly subtle difficulties are created, and, in language-test items especially, variations may be contorted to the point where the validity of the whole exercise may be questioned.

In simple terms, the question to which the teacher, the curriculum consultant, or the trustee is likely to want an answer is, "How many or what proportion of students by the end of Grade 8 can write a passage of 150 words in complete sentences?". As soon as one turns this question into a statement, one has a plausible objective: "The student by the end of Grade 8 can write a passage of at least X words with a frequency of no more than Y 'no-sentence' errors." In this form it is a criterion for the individual student. Note that it is either *met* or it is *not met*.

For a class, a grade or program level, or a system, the objective might be stated as: "By the end of Grade 8 at least X per cent of the students can write a passage of Y words with a frequency of no greater than Z 'no-sentence' errors."

This very simple (and popular) objective illustrates three features of criterion-referenced testing:

1. A criterion-referenced test does not have to be a multiple-choice test. Obviously, a test for this objective requires that the student write under reasonably controlled conditions.

2. The measures of performance are fairly direct. (This one happens to be easily "countable", though many important aspects of English are not.) They should also be readily understandable and capable of being repeated in the future with other students or with the same student to measure change in performance.

3. The criteria chosen are quite arbitrary. For example, what percentage of students, by the end of Grade 8, "should" be making no "no-sentence" errors, no grammatical howlers, etc.? Many factors must be considered in establishing reasonable performance criteria, whether the criterion in question is "number of 'no-sentence' errors" or "frequency of oral response" in the classroom; for example, what if a particular school population includes a higher-than-average proportion of students from lower socio-economic backgrounds, or of students for whom English is a second language, or of students of a different cultural background?

It follows that, though we do not need a norm-referenced test (i.e., one yielding a wide spread of scores), criteria should have their basis in what can reasonably be expected of most children at a given age or stage of education. Wherever possible, this sort of information should be used in specifying objectives and designing programs. However, teachers can also obtain

[1]W. J. Popham, *An Evaluation Guidebook* (Los Angeles: Instructional Objectives Exchange, 1972), p. 32.

a fairly accurate picture of general student ability by counting or judging performance for a wide range of criteria from an adequate sample of student writing. Thus, some loose norms can be established as a starting point, the principle being that the criteria established should be such that most students can be expected to meet most of them.

In order to determine the performance level across a system on a number of criteria, it is not necessary to test every student. A carefully drawn random sample will do very well; there is no need to identify student, teacher, or school if what is wanted is an overview of performance on a number of specific objectives pertinent to the program.

Besides directly observing student writing, a teacher or group of teachers should design short tests (which can be scored by students) to measure achievement in a variety of areas of English, the aim being – besides the testing of a particular instructional unit – to build a bank of appropriate test items, some of which may be reserved for system monitoring and others of which may be employed at appropriate times with a class or with individual students to determine specific problems and needs.

A short test to detect ''no-sentence'' errors, for example, might consist of the following: the teacher selects a paragraph written by a student working at the grade level being tested and removes all punctuation marks and capital letters at the beginning of sentences. The teacher then duplicates the passage and has all the students provide end punctuation.

As a second example, assume that the problem or lesson has to do with connectives – their recognition and effective use. To test for recognition, the teacher might select a passage (either from the students' writing or from textual material that students at this level are required to read) or create one, duplicate it, and have the students circle the connectives. Alternatively, the teacher could eliminate the connectives and have the students fill in what they believe to be connectives appropriate to the logic of the passage. Multiple-choice variations can also, with more difficulty, be developed; here, a selection of less and more effective connectives can be provided for each blank.

Teacher exchanges of short tests will lead to the improvement of the tests. Refined versions can then be kept on file for future use. Finally, a selection of test items can be made for monitoring student achievement across the system. This process will proceed most effectively if there is teacher support and sharing.

Extension to the System

With testing strategies keyed to important program objectives, a kind of ''master plan'' can be developed to ensure that most student performance can be measured in terms of course objectives. A matrix[2] can then be established, with test items or strategies keyed to it so that a testing plan can be constructed that samples all or most program objectives.

For testing on a system basis, particularly in language and reading, different students may be given different (and relatively few) items, so that performance across the system is measured for all objectives, though the testing time is short. For this type of testing, it is important that items be well refined and pretested, and that most items be scorable by computer (not so much because scoring itself is onerous, but because the management and interpretation of the data is complicated).

Besides keeping the testing time short for the individual student, the advantages of this type of system-wide testing include the following:

- The items have been designed by the teachers.
- They have been carefully scrutinized for content validity.
- They represent many program objectives, all considered important by the teachers of English in the system.
- The test combines flexibility (sections or items can be deleted or added) and stability (changes in performance relative to particular objectives can be noted over time) and makes possible conclusions about the need for, or effect of, changes in course content or teaching strategy.
- The testing procedure is not threatening. Students, teachers, and schools need not be identified.

The most difficult task connected with this kind of testing lies at the beginning: the task of selecting particular objectives from the guideline in planning courses of study and of refining these objectives to the point where performance can be evaluated. This can be done gradually and a balance can be maintained among the objectives. The books in the general bibliography to this document that deal specifically with the elaboration of objectives can be consulted. The more carefully objectives are established, the simpler and fairer the evaluation of individual student performance and of the level of performance across the system will be.

It is helpful, in the development of effective, useful evaluation strategies in English, for teachers, both within a school and within a system as a whole, to share resources. The following statement offers some suggestions in this regard:

> It is recommended that:
> 2. There be developed a bank of specific objectives and corresponding test items and instruments (we would add ''strategies'') based on curriculum objectives that would be available to teachers for their use in the diagnostic testing of students.
>
> (We believe that the appropriate beginning for specification of objectives and content-valid tests and strategies must be with practising teachers of English.)
>
> 3. The Ministry provide assistance to school boards which may wish to develop local testing programs. Where such assistance leads to the production of new items or instruments, the sharing of these with other jurisdictions in the province should be expected.[3]

The Ontario Assessment Instrument Pool, currently being developed, will make available to teachers

[2]See, for example, B. S. Bloom et al., *Handbook of Formative and Summative Evaluation of Student Learning* (New York: McGraw-Hill, 1971).

[3]*Report of the Work Group on Evaluation and Reporting,* p. 16.

specific instruments and strategies related to objectives outlined in the Intermediate Division English guideline.

Monitoring

Many of the recommendations in the preceding section are also suggested by the 1975 Bullock Report, *A Language for Life,* which analyses the teaching of English in England and Wales. Chapter 3, "Monitoring", which proposes an annual national monitoring of standards, particularly in language study, reading, and writing, is particularly worth examining.

The British report proposes, as its strongest recommendation, the sampling of actual student writing holistically scored, with a number of essays from previous years included as a means of estimating the advance or decline in standards. Important points are also made concerning measures of syntactic maturity and the relation of writing style to purpose.

The item pool, referenced to a wide range of specific objectives, is also recommended for language and reading.

> In the monitoring of reading and writing we recommend a new style of assessment which will allow for an extensive coverage of attainments without imposing a heavy testing load on individual pupils. The principle we suggest is the sharing of a selection of assessment tasks between a number of groups of pupils. At any one phase of assessment each group attempts a different set of exercises or items from that of every other group. The performance of the population is thus estimated by the performance of the separate groups taken together. The levels of attainment in a single test will be represented as the mean score obtained by each sample of pupils.[4]

The Bullock Report points out other advantages of "item-pooling", whether one considers the practice on a large scale or within a system or school:

> A signal advantage of question-pooling is that it offers a degree of flexibility the single test can never provide. When the monitoring surveys are in train new exercises can be tried out alongside the calibrated items and thus "chained in" at the appropriate point. Out of date material or examples found to be unsatisfactory can be discarded, while the repeated inclusion of an item will provide data which can be used to improve the accuracy of calibration. The major benefit of this flexibility is that it will be possible for the monitoring system to keep abreast of changes in the use of language and in teaching emphases in school.[5]

Because of the nature of British education, the committee recommends testing of eleven- and fifteen-year-olds. In order that monitoring should not be a massive or traumatic undertaking, the committee recommends a sampling in each term of only sixteen secondary or thirty-two Primary schools on each occasion.

> 1,600 pupils would be required at one time and eight of the tests from the pool would be divided among them, so that each test was completed by 200 children. By covering eight features of attainment in this manner, it would be possible to gain a great deal of information without increasing the demand upon any one school or pupil. As a general rule a school would be selected only once in several decades.[6]

Although provincially we may be at some distance from the systematic monitoring recommended in the Bullock Report, the strategies recommended do seem feasible for a school or a system to undertake in order to achieve a reasonable measure of student achievement in English.

Reporting

However comprehensive evaluation schemes may be, their validity and reliability will only be accepted if their audience clearly understands what is being evaluated. The responsibility of the teacher in reporting achievement is to make the aims and objectives outlined in the Intermediate Division English guideline (pages 5 and 10-13) an integral, visible part of the reporting system.

In order to incorporate the aims and objectives in English programs in the Intermediate Division, specific criteria should be elaborated for the various aspects of the program and reflected in the reporting scheme. The following two Ministry documents will prove useful in this regard:

- *Evaluation of Student Achievement: A Resource Guide for Teachers, 1976;* and
- *Report to the Minister of Education, the Hon. Thomas L. Wells, of the Work Group on Evaluation and Reporting, 1977.*

The section entitled "Reporting" from the *Report on Evaluation and Reporting* contains some suggestions that are particularly relevant to English. These are paraphrased here:

- The reporting should touch on both areas of strength and areas needing improvement.
- The student should have help in developing the skills of self-evaluation.
- There should be frequent, informal reports to the student.
- In reporting to parents, an explanation of program objectives should constitute part of the reporting process.
- There should be no doubt left in the parents' minds as to the state of the student's progress.
- Good in-school reporting requires the maintenance of student records, containing current and complete information.

[4]Committee of Inquiry into Reading and the Use of English, *A Language for Life* (London: Her Majesty's Stationery Office, 1975), p. 41.

[5]Ibid., p. 42.

[6]Ibid., p. 42.

Chapter IV, "Communicating Student Achievement", in *Evaluation of Student Achievement* explores the characteristics of effective reporting techniques, several of which have particular relevance to reporting in English. There are cautions about just how much (and how often) information should be presented, and about the relationship between effective reporting and the need of the recipients to understand the program's goals, standards, and objectives. These documents promote a reporting process characterized by clarity and specificity. The adoption of such a process will provide everyone concerned with sufficient information about the actual achievement of students and the effectiveness of programs.

Bibliography

General

Bloom, B. S.; Hastings, J. I.; and Madaus, G. F. *Handbook on Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill, 1971.

Besides specific chapters concerning English (language arts, literature, and writing), this is a most useful general book on formative and summative evaluation, test design and development, and range of objectives in both cognitive and affective areas of learning.

Buros, O. K., ed. *English Tests and Reviews*. Highland Park, N. J.: Gryphon, 1975.

This is an authoritative analysis and review of many of the standardized tests of English. It could serve as an important resource for superintendents and consultants or others involved in decisions concerning test selection.

Cooper, C. R., and Purves, A. C. *A Guide To Evaluation*. New York: Ginn, 1974.

This compendium of strategies for measuring comprehension, response to literature, group and individual performance, self- and peer evaluation, and attitudes is recommended.

Department of Education and Science, United Kingdom. *A Language for Life*. Report of the Committee of Inquiry . . . under the Chairmanship of Sir Alan Bullock. London: Her Majesty's Stationery Office,1975. Available through Pendragon House, Toronto.

Known as the "Bullock Report", this is perhaps the most thorough, interesting, and comprehensive report ever made on the teaching, learning, and evaluating of first-language learning. The entire report is valuable in providing insights on curriculum and language learning. The opening chapters on attitudes and standards are most pertinent to English evaluation. The report contains, as well, a valuable chapter which reviews the efforts to measure reading-achievement levels over a period of time in England and Wales and contains instructive recommendations on test use and interpretation. The brief chapter on monitoring provides excellent recommendations for the effective measurement of literacy on a system-, province-, or nation-wide basis.

Mellon, J. C. *National Assessment and the Teaching of English*. Champaign, Ill.: National Council of Teachers of English, 1975.

This is an analysis and report of the first round of the National Assessment of Educational Progress in the United States. Besides examining the implications of national testing, it provides a close look at testing in writing, reading, and literature.

National Council of Teachers of English. *Common Sense and Testing in English*. Report of the Task Force on Measurement and Evaluation in the Study of English. Champaign, Ill.: NCTE, 1975.

This report focusses on norm-referenced tests and the legitimate as opposed to the frequently illegitimate use of test scores. It contains important notes on "culture-fair" testing and advice to teachers on test selection. At $1.00, it is an indispensable resource, and eminently practical.

Popham, W. J. *An Evaluation Guidebook.* Los Angeles: Instructional Objectives Exchange, 1972.

This is a set of guidelines for the educational evaluator. It is practical, blunt, and perhaps too strong on "behavioural objectives" for the taste of most teachers of English. Nevertheless, it is a useful antidote to generalized evaluation. It is especially strong on the distinction between criterion-referenced and norm-referenced testing and the advantages of the former.

Stibbs, Andrew. *Assessing Children's Language: Guidelines for Teachers.* London: Ward Lock Educational and The National Association for the Teaching of English, 1979.

This inexpensive little volume is concise, clearly written, and practical in its approach. It examines traditional ways of assessing students' language and suggests sound and useful alternatives.

Journals

The English Journal, Elementary English, and *Research in the Teaching of English* are all published by the National Council of Teachers of English (NCTE). These and other NCTE publications, as well as numerous British publications on English teaching, are available from:
National Council of Teachers of English
1111 Kenyon Road
Urbana, Illinois 61801

Highway One and *The English Quarterly* are published by the Canadian Council of Teachers of English, 237 Yonge Blvd., Toronto, Ontario M5M 3J1.

Objectives in English

The art of stating with clarity and reasonable comprehensiveness a range of objectives appropriate to English will be greatly improved by the use of the following resources:

Lazarus, A., and Knudson, R. *Selected Objectives for the English Language Arts, Grades 7-12.* Boston: Houghton Mifflin, 1967.

Although it is frustrating to be forced to recommend a book currently out of print, this volume is so helpful as to justify its inclusion in this bibliography. There are chapters (each with an excellent bibliography) on every facet of the study of English. There is also a valuable appendix on language study for teachers. Each section makes distinctions among attitudes, understandings, skills, and habits. Objectives are generally not defined in strict "behavioural" form.

National Assessment of Educational Progress. *Reading Objectives – Second Assessment.* Denver: NAEP, 1974. Available, along with similar booklets and reports on other aspects of NAEP, including writing and literature, from the NAEP, Lincoln Tower, 1860 Lincoln St., Denver, Colo. 80203.

Based on one reasonably successful round of evaluation, this twenty-page booklet is a revision and restatement of reading objectives considered suitable for testing. It is particularly notable for its elaboration of "reading comprehension".

Tri-University Project on Behavioural Objectives in English. *Representative Performance Objectives for High School English.* New York: Ronald Press, 1971.

The opening chapter of this book contains helpful advice both on the use of the book and on efforts to elaborate criteria for performance in English. The distinction it makes between "performance" objectives and "enabling" objectives is interesting and will help teachers to distinguish between primary and secondary objectives. *Note: As of 1976 this book is in the public domain and may be reproduced without charge or penalty.*

Reading

Farr, R. *Reading: What Can Be Measured?* Newark, Del.: International Reading Association, 1969.

This is a very important resource for understanding the limits of testing in reading and the specific weaknesses of, and misrepresentations that result from, current standardized tests of reading. It contains a vigorous attack on existing tests and testing strategies for reading.

Ruddell, R. B., ed. *Accountability and Reading Instruction: Issues.* Champaign, Ill.: National Council of Teachers of English, 1973.

Particular reference is made to Kenneth Goodman's critique of current reading-test practices, with notes concerning statistical fallacies and the lack of an articulated theory of reading.

Venezky, R. L. *Testing in Reading.* Champaign, Ill.: National Council of Teachers of English – Educational Resources Information Center (ERIC), 1974.

This is a brief booklet listing ten canons for both instructor and administrator, which cover testing purposes and test selection. Each canon is elaborated. Though the title suggests that the booklet limits itself to a consideration of reading, each principle is appropriate to decision-making about tests and the use of test scores in any subject.

Writing

Britton, J. N.; Martin, N. C.; and Rosen, H. *Multiple Marking of English Compositions.* Schools Council Examination Bulletin no. 12. London: Her Majesty's Stationery Office, 1966.

This bulletin describes a highly reliable method of scoring essays by using teams of rapid-impression markers, together with strategies for the evaluation of a large sample of each student's writing over the year. Holistic scoring and other features of this research project have been used in Ontario in the 1976 Interface study and in the scoring of the writing of students in Grades 7 and 8 with eighteen school boards (1976-77).

Similar features appear in the American National Assessment of the writing of nine-, thirteen-, and seventeen-year-olds. When adapted, the method is suitable for use in evaluating writing performance in a school or system.

Britton, J. N.; Burgess, T.; Martin, N. C.; McLeod, A.; and Rosen, H. *The Development of Writing Abilities (11-18).* London: Macmillan, 1975.

This book carefully examines modes of writing in terms of intended function and audience, their appropriateness to particular age levels, and their frequency of assignment in English and other parts of the school program. It is an important resource both for the reconsideration of writing programs in schools and the establishment of criteria for writing performance. It will be of greater benefit to whole departments of English or to a family of schools than to individual teachers working on their own.

Cooper, C. R. "Measuring Growth in Writing". *English Journal,* 64 (March 1975), 111-20.

The use of this review of research in the areas of language and writing development has already been discussed in the chapters on language and writing. A number of strategies for measuring writing achievement are presented.

Cooper, C. R., and Odell, Lee, eds. *Evaluating Writing: Describing, Measuring, Judging.* Champaign, Ill.: National Council of Teachers of English, 1977.

This publication includes excellent essays on practices in holistic and other forms of scoring, including reference to the interesting experimental work in judging essays that was conducted in the American National Assessment. It also contains an approach to self- and peer evaluation of writing, and an essay concerning the development of syntactic structures as students mature.

Diederich, P. B. *Measuring Growth in English.* Champaign, Ill.: National Council of Teachers of English, 1974.

This book presents an interesting approach to reading evaluation, although for the most part it deals with the evaluation of writing. The approach it takes is direct and reliable, with recommendations for the operation of a program of regular evaluation of writing development across grades and schools. The book was written for English teachers who feel rather insecure with the technical jargon of measurement.

Evans, P.; Brown, P.; and Marsh, M. *Criteria for the Evaluation of Student Writing, Grades 7 and 8: A Handbook.* Toronto: Ontario Institute for Studies in Education, 1976.

This publication contains both criteria for the evaluation of a number of modes of writing appropriate especially to the Intermediate Division and additional general criteria applicable to many modes. Criteria were established by teachers of Grades 7 and 8 who participated in a writing-evaluation project. Eighteen Ontario school systems are represented. The handbook would be useful either within the classroom or as a guide to schools or systems elaborating their own criteria for Grade 7 and up.

Follman, J. C., and Anderson, J. "An Investigation of the Reliability of Five Procedures for Grading English Themes". *Research in the Teaching of English* 1 (1967): 190-200.

There are many journal articles reporting or recommending strategies for evaluating writing. This very clearly written article reviews with care five of the more useful procedures, their strengths and their weaknesses. Consequently, it is an excellent introduction for teachers attempting to come to grips with the problem of reliability or consistency in evaluating writing.

Traub, Ross; Wolfe, Richard; Wolfe, Carla; Evans, Peter; and Russell, H. H. *Secondary-Postsecondary Interface Project II: Nature of Students.* 2 vols. Toronto: Ontario Ministries of Education and of Colleges and Universities, 1976.

Pertinent sections describing the measurement of writing achievement are found in volume 2, pages 38-118. Instructions to holistic scorers and to error-counters are also included. Variations on this method may be employed for rapid impression-scoring of student essays at various grade levels and for close analysis of the characteristics of writing to be found in a sample.

Notes

THE LIBRARY

The Ontario Institute

for Studies in Education

Toronto, Canada

9048
ISBN 0-7743-4050-9